The second law and fluctuations of work: The case against quantum fluctuation theorems.

A.E. Allahverdyan[1,2)] and Th.M. Nieuwenhuizen[1)]
[1)] *Institute for Theoretical Physics, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands and*
[2)] *Yerevan Physics Institute, Alikhanian Brothers St. 2, Yerevan 375036, Armenia*

We study how Thomson's formulation of the second law: no work is extracted from an equilibrium ensemble by a cyclic process, emerges in the quantum situation through the averaging over fluctuations of work. The latter concept is carefully defined for an ensemble of quantum systems interacting with macroscopic sources of work. The approach is based on first splitting a mixed quantum ensemble into pure subensembles, which according to quantum mechanics are maximally complete and irreducible. The splitting is done by filtering the outcomes of a measurement process.

A critical review is given of two other approaches to fluctuations of work proposed in the literature. It is shown that in contrast to those ones, the present definition *i)* is consistent with the physical meaning of the concept of work as mechanical energy lost by the macroscopic sources, or, equivalently, as the average energy acquired by the ensemble; *ii)* applies to an arbitrary non-equilibrium state. There is no direct generalization of the classical work-fluctuation theorem to the proper quantum domain. This implies some non-classical scenarios for the emergence of the second law.

PACS numbers: PACS: 05.30.-d, 05.70.Ln

I. INTRODUCTION.

The second law was deduced in XIX'th century, and formulated for a single closed system, in a way resembling the laws of mechanics [1, 2, 3, 4, 5, 6]. It was, however, already the insight of Maxwell [7, 8] and Gibbs [9] that this law has in fact a statistical character, and refers to averages over an ensemble of identically prepared systems, rather than to a single system. This viewpoint became widely accepted since the beginning of XX century, where first robust observations of fluctuations were made [1]. Together with theoretical works of Boltzmann in kinetic theory of gases and of Smoluchowski, Fokker, Planck and Einstein in physics of brownian motion, they formed a consistent picture of the second law as emerging from microphysics through averaging over fluctuations. A detailed summary of this activity is presented in the book by Epstein [1], while Tolman [2] discusses theoretical aspects of the situation. Since then, the statistical understanding of the second law entered several modern books of statistical physics and thermodynamics [3, 4]. The current perspectives on the classical and quantum brownian motion in the context of the second law can be found in Refs. [10, 11, 12].

At the end of 1970's several groups independently gave a derivation of Thomson's formulation of the second law [13, 14, 15, 16, 17, 18, 19]: *no work can be extracted from initially canonical equilibrium system by means of a cyclic thermally isolated process*, starting directly from quantum or classical Hamiltonian equations of motion. The very possibility of getting such a straightforward thermodynamical result directly from equations of motion is due to the fact that work is a transparent quantity unambiguously defined both in and out of equilibrium for any (quantum or classical) system interacting with external macroscopic work sources [2]. As the main consequence, Thomson's formulation is the only one which is valid both for finite and infinite systems which do start in equilibrium, but can be driven arbitrary far from it by external sources (see [20] and sections II C and VII for more details).

The standard understanding of the second law and fluctuations is based on Einstein's formula relating entropy to the probability of a fluctuation around equilibrium [1, 2, 3, 5]. This suffices for the purposes of near-equilibrium thermodynamics of macroscopic bodies, in particular, because all the formulations of the second law are equivalent for them and entropy is defined unambiguously. In the more general case of finite systems and/or systems driven strongly out of their initial equilibrium, relations between the second law and fluctuations ought to be studied anew for each meaningful formulation of the law in separate. The conclusions may differ from case to case and from formulation to formulation [12, 20].

The purpose of the present paper is to understand how Thomson's formulation of the second law in the quantum situation emerges through the averaging over fluctuations of the equilibrium ensemble. More specifically, if the (average) work done on the initially equilibrium ensemble during a cyclic process is always non-negative, what are

[1] It was thus rather surprising to see recent claims on "violations of the second law" [23] or "transient violations of the second law" [24] due to fluctuations; see in this context our comment [25].

[2] These features of work are in contrast to those of entropy, whose meaning is too closely tied to equilibrium states of macroscopic bodies.

fluctuations of this work, and how do they behave? There are definite answers to these questions in the classical situation: the definition of fluctuations of work is straightforward, and a model-independent information on them is given by an equality first derived by Bochkov and Kuzovlev in 1978 [13] (BK equality). Later on, this equality, sometimes also called work-fluctuation theorem, was extended to non-cyclic processes [21], and has undergone various generalizations [3]. The basic messages of the classical BK equality are recollected and reviewed in section II E.

While all these developments concern the classical situation, a number of recent works is devoted to quantum extensions of BK equality [27, 28, 29, 30, 31, 32]. The first definition of fluctuations of work in the quantum situation and a quantum extension of BK equality was actually proposed by Bochkov and Kuzovlev themselves [13, 15]. It is based on constructing an operator in the Heisenberg representation, associating it with an operator of work, and thus treating work as an ordinary quantum mechanical observable pertaining to the system and not to the work-source.

Another extension was initiated by Kurchan [28], based on two-time measurements of energy. As implied by the standard treatment of quantum measurements, this second approach is closely tied to Schrödinger representation.

There are therefore two different approaches to the definition of fluctuations of work and to quantum extensions of BK equality; both of them attracted attention recently [27, 29, 30, 31, 32], and are reviewed below in section VI. The fact that in the quantum situation these approaches are different is calling for attention. The difference in viewpoints is not completely unexpected, since the work as it appears in statistical thermodynamics [2, 3, 4, 5] is an essentially classical quantity (mechanical energy transferred from a classical source of work).

Our objective is to propose a third definition of fluctuations of work, which is motivated by the fundamental physics of quantum (sub)ensembles. The definition is guided by the following observation. Since the usual work is now presented as an average of a random quantity — for the moment we leave unspecified whether this is a random classical quantity or an operator — it is natural to require the following two conditions on its (fluctuating) realizations and on its average:

- Once the average work is unambiguously defined for any quantum system starting in an arbitrary initial state and interacting with a macroscopic source of work, the same should hold for fluctuations of work. In particular, one *cannot* restrict the general definition to (initially) equilibrium states of the system. This is relevant, since one of the chapters in statistical thermodynamics deals with work extraction from non-equilibrium systems [3, 40], and one should, of course, be able to define fluctuations of work in this most general situation.

- Realizations of the random quantity work should have the same physical meaning of mechanical (high-graded) energy as the usual (average) work. In particular, if one happens to extract some work from a single realization, one should be able — at least in principle — to use precisely this amount for the standard purposes, e.g. for driving motors. (Basic features of work are recalled below in section II B.)

Both these conditions are satisfied by the classical definition, and to our opinion without them the very program of studying the emergence of the second law in the quantum situation becomes ill defined.

It appears to the present authors that, as we discuss below in section VI, neither of the existing two quantum approaches —in the way they stand presently— can be viewed as providing a proper definition of fluctuations of work in the quantum situation. Both approaches fail out of equilibrium (no first condition), while even for an initially equilibrium state it is not clear that the second condition is satisfied. Neither these points were discussed in papers which support those definitions; see e.g., [13, 15, 27, 28, 30, 31, 32].

These are the reasons to introduce in the present paper a third approach to quantum fluctuations of work, which will satisfy the above two conditions. Our approach starts with explicitly respecting the first condition, that is, always defining realizations of (the random quantity) work as an average energy given off by the macroscopic source of work. If the corresponding ensemble of physical systems already consists of subensembles, non-trivial realizations can be defined via the average energy exchange of each subensemble with the source. (As with any exchange process, this is operationally characterized by measurements at two different times.) For a classical ensemble each single member completely characterizes a subensemble, and the classical definition of fluctuations of work follows naturally. In contrast, a quantum equilibrium state is described by a homogeneous quantum ensemble, the gibbsian, which by itself does not consist of subensembles. This prevents us to proceed as such. First, an inhomogeneous quantum ensemble has to be prepared from this homogeneous one, a task accomplished via a selective quantum measurement. The obtained structure of subensembles does depend on the type of measurement, and as a consequence the resulting fluctuations of work in the quantum situation appear to be context-dependent (contextual). Still some relevant characteristics of these fluctuations can be context-independent, as seen below.

[3] A rather complete account of various generalizations of the classical work-fluctuation theorem, as well as its relation with other fluctuation theorems, e.g., those describing entropy production, is given in Refs. [22, 26]. Local versions of the fluctuation theorems are also discussed there.

As the main result, the second law in Thomson's formulation — whose statement reads in the same way both in quantum and classics — has in those two situations rather different scenario of emergence. The basic qualitative difference is that in contrast to classics, the fluctuations of work in the proper quantum situation are not controlled by any direct analog of BK equality. More specifically, in classics the structure of work as a random quantity is such that there have to be realizations which provide work (i.e., which are active). In the quantum case, however, there need not be any active realization.

In this paper we have taken the simplest approach which allows to study Thomson's formulation of the second law and fluctuations of work, that is, we consider a finite quantum or classical system interacting with external sources of work. The restriction to finite — though possibly large — systems is at any rate natural for studying fluctuations. The explicit presence of thermal baths, as well as details about the thermodynamical limit for the system are left for future studies.

The paper is intended to be self-consistent and is organized as follows. In section II we recall the definition of fluctuations of work in the classical situation and review the BK equality and its consequences. In section III we present the definition of fluctuations of work in the quantum situation. The dispersion of work is studied in section IV. In section V we show that fluctuations of work in the quantum situation are not controlled by any direct analog of the classical BK equality. An anti-classical scenario for the emergence of the second law in Thomson's formulation is described in section IV. In section VI we make comparison with the two known approaches on fluctuations of work in the quantum situation. These approaches offer different extensions of classical BK equality. We do not intend to imply that these approaches do not have a physical meaning or that they cannot be useful for their own sake. We only state that —in the way they stand presently— they do not describe fluctuations of work in the proper quantum situation. We close with a discussion. Some details are worked out in appendices.

II. CLASSICAL FLUCTUATIONS OF WORK AND BK EQUALITY.

A. The setup.

Consider an ensemble $\mathcal{E}$ of identical classical systems $\mathcal{S}$ which are thermally isolated [3, 4]: they move according to their own dynamics and interact with an external macroscopic work source $\mathcal{W}$. This interaction is described via time-dependence of some parameters $R(t) = \{R_1(t), R_2(t), ...\}$ of the system's Hamiltonian $H(t) = H\{R(t)\}$; see Refs. [3, 4].

The parameters move along a certain trajectory $R(t)$ which at some initial time $t = 0$ starts from $R(0)$, and ends at $R(\tau)$ at the final time $t = \tau$. Cyclic thermally isolated processes are defined by $R(0) = R(\tau)$ and thus

$$H\{R(\tau)\} = H\{R(0)\} \equiv H. \tag{1}$$

At the initial time the ensemble is in equilibrium, that is, the common probability distribution $\mathcal{P}(x, p; t = 0) \equiv \mathcal{P}(x, p)$ of all its canonically conjugated coordinates $x = (x_1, ..., x_n)$ and momenta $p = (p_1, ..., p_n)$ is given by the Gibbs distribution with the initial Hamiltonian $H(x, p)$ and temperature $T = 1/\beta \geq 0$:

$$\mathcal{P}(x, p) = \frac{e^{-\beta H(x,p)}}{Z}, \quad Z = \int \mathrm{d}x\, \mathrm{d}p\, e^{-\beta H(x,p)}. \tag{2}$$

This equilibrium distribution can be prepared by means of a thermal bath coupled with the system $\mathcal{S}$ for $t < 0$. It is assumed that for times $0 \leq t \leq \tau$ the system $\mathcal{S}$ is decoupled from the bath —an alternative assumption would be that its coupling to the bath is so weak that it can be neglected in the considered time-interval— and the evolution of the ensemble is described by the Liouville equation for $\mathcal{P}(x, p; t)$:

$$\partial_t \mathcal{P}(x, p; t) = \frac{\partial H(x, p, t)}{\partial p} \frac{\partial \mathcal{P}(x, p; t)}{\partial x} - \frac{\partial H(x, p, t)}{\partial x} \frac{\partial \mathcal{P}(x, p; t)}{\partial p}. \tag{3}$$

B. Work

In statistical thermodynamics there are two alternative definitions of work [2, 3, 4, 6, 33]. Both are necessary for the proper understanding of its physical meaning [4, 33, 42]. The first reads

- The work W is the average energy gained by $\mathcal{S}$ during a thermally isolated system-work-source interaction with $\mathcal{W}$ [3, 4]:

$$W = \int \mathrm{d}x\, \mathrm{d}p\, [\, \mathcal{P}(x, p; \tau)\, H(x, p; \tau) - \mathcal{P}(x, p)\, H(x, p; 0)\,]. \tag{4}$$

Due to conservation of energy, W is equal to the average energy lost by the work-source $\mathcal{W}$. This definition was (implicitly) proposed by Caratheodory [33]. A concise history of various definitions of work is given in [38], while various perspectives of work in classical mechanics are reviewed in [39].

For cyclic processes Eq. (4) takes a simpler form

$$W = \int \mathrm{d}x\, \mathrm{d}p\, [\mathcal{P}(x, p; \tau) - \mathcal{P}(x, p)]\, H(x, p; 0). \tag{5}$$

There is a second, alternative definition going back to Gibbs and Planck [33, 38]:

- The minus work $-W$ is the energy transferred to the work-source $\mathcal{W}$. Its distinguishing feature with respect to other forms of energy is that it can, in principle, be transferred with 100 % efficiency to other work sources via interactions of the system-work-source type. In particular, it can be retransferred to collective degrees of freedom that perform *classical deterministic* motion generated by a suitable Hamiltonian. These degrees of freedom are thus purely mechanical and serve as prototypes of macroscopic mechanical devices (such as a motor, piston, turbine, etc.). For them the differential work can be calculated in the usual way of ordinary mechanics, that is, multiplying the external force by the corresponding displacement [4].

Both these definitions of work are expected to be equivalent at least for sufficiently ideal work sources [4, 33, 42].

Yet we mention for completeness another, equivalent formula for the work W; the integral of the rate of energy change:

$$W = \int_0^\tau \mathrm{d}t \int \mathrm{d}x\,\mathrm{d}p\,\mathcal{P}(x,p;t)\,\frac{\partial H(x,p;t)}{\partial t}. \qquad (6)$$

To get from here to Eq. (5) one performs integration by parts, uses the standard boundary conditions, that is $\mathcal{P}(x,p;t)$ decays for $x \to \pm\infty$ or $p \to \pm\infty$ and employs Eq. (3). This formula for W is more general and can be applied to processes that involve explicit thermal baths.

C. Fluctuations of work

Though the ensemble $\mathcal{E}$ is described by the probability distribution $\mathcal{P}(x,p)$, each single system $\mathcal{S}$ from this ensemble has at a given moment of time explicit values for all its dynamical variables. These values may vary from one single system to another due to the distribution of initial conditions.

Each single member of the ensemble is then coupled to the external source of work that realizes on it a unique thermally isolated process (the same for all members). In other words, the same parameters $R(t)$ of the Hamiltonian are varied in the same way for each member. The motion of the single system is described by Eq. (3) with now $\mathcal{P}(x,p;t)$ being a product of two delta-functions $\delta(x - x(t))\,\delta(p - p(t))$, which are probability densities concentrated at the solutions of the canonical equations of motion:

$$\dot{p} = -\partial_x H(x,p;t), \qquad \dot{x} = \partial_p H(x,p;t). \qquad (7)$$

The trajectories generated by (7), together with their initial conditions distributed according to Eq. (2), serve as *realizations* of the random process given by Eq. (3).

The work $w(x,p)$ exchanged in each thermally isolated process can then be calculated consistently with Eq. (5):

$$w(x,p) = H(x(\tau),p(\tau);\tau) - H(x,p) \qquad (8)$$
$$= H(x(\tau),p(\tau)) - H(x,p), \qquad (9)$$

where $H(x(\tau),p(\tau);\tau)$ is the value of the Hamiltonian on the trajectory that started at $t=0$ from (x,p), with $x(\tau)$ and $p(\tau)$ being the corresponding solutions of (7). This work can be observed as the energy decrease of the mechanical degree of freedom of the macroscopic work-source, or alternatively via energy increase of the system $\mathcal{S}$. In this latter scenario the energy of $\mathcal{S}$ has to be measured two times, at the moments $t=0$ and $t=\tau$.

The so defined work $w(x,p)$ for a single system is a random quantity, since it varies from one single system to another. It can be positive or negative. Its probability distribution $P(w)$ is determined by $\mathcal{P}(x,p)$, since this is the probability by which each single system enters in the ensemble:

$$P(w) = \int \mathrm{d}x\,\mathrm{d}p\,\mathcal{P}(x,p)\delta(w - w(x,p)). \qquad (10)$$

There being used no special features of the initial equilibrium distribution function, the same definition for the work in a single realization can be given for any initial ensemble.

It is seen that the two desired conditions for fluctuations of work formulated in section I are naturally satisfied: the initial distribution may be arbitrary and "work for a single realization" has the same physical meaning as average work.

D. Derivation of BK equality.

One now derives BK equality in the classical situation for a closed cycle, [13, 15, 21]:

$$\langle e^{-\beta w}\rangle \equiv \int \mathrm{d}w\,P(w)\,e^{-\beta w} \qquad (11)$$
$$= \int \mathrm{d}x\,\mathrm{d}p\,\mathcal{P}(x,p;0)\,e^{-\beta w(x,p)}$$
$$= \frac{1}{Z(0)}\int \mathrm{d}x\,\mathrm{d}p\,e^{-\beta H(x,p)-\beta w(x,p)}$$
$$= \frac{1}{Z(0)}\int \mathrm{d}x\,\mathrm{d}p\,e^{-\beta H(x(\tau),p(\tau);\tau}$$
$$= \frac{1}{Z(0)}\int \mathrm{d}x(\tau)\,\mathrm{d}p(\tau)\,e^{-\beta H(x(\tau),p(\tau);\tau)}$$
$$= \frac{Z(\tau)}{Z(0)} = 1, \qquad (12)$$

where we used Liouville' theorem $\mathrm{d}x\,\mathrm{d}p = \mathrm{d}x(\tau)\,\mathrm{d}p(\tau)$ and Eqs. (2, 9, 10). The last equality in (12) is due to the assumed cyclic feature of the process.

E. Qualitative messages of the BK equality.

The BK equality is by itself an exact mathematical relation. Several important qualitative results can be deduced from it:

a. The second law. As the exponential function is convex, one gets directly $1 = \langle e^{-\beta w}\rangle \geq e^{-\beta\langle w\rangle}$, and then $W = \langle w\rangle \geq 0$, which is the statement of the second law in Thomson's formulation: no work can be extracted from an equilibrium system by means of a cyclic process. This formulation of the second law is well-known and has an independent and more general derivation both in the classical and the quantum situation [13, 16, 17, 18, 19, 35].

b. Active realizations. To satisfy $1 = \langle e^{-\beta w}\rangle$ directly leads to the following observation: for any cyclic thermally isolated process there are realizations which are active, that is, for which work is extracted after the process: $w(x,p) < 0$. The relative weight of such active realizations can be estimated via the Cauchy inequality:

$$1 = \left(\int \mathrm{d}x\,\mathrm{d}p\,\sqrt{\mathcal{P}(x,p)}\,\sqrt{\mathcal{P}(x,p)}\,e^{-\beta w(x,p)}\right)^2 \leq \int \mathrm{d}x\,\mathrm{d}p\,\mathcal{P}(x,p)\int \mathrm{d}x\,\mathrm{d}p\,\mathcal{P}(x,p)\,e^{-2\beta w(x,p)}, \quad (13)$$

which can be written as

$$\langle e^{-2\beta w}\rangle \geq 1. \quad (14)$$

A stronger relation can be obtained with help of a generalization of the Cauchy inequality, described in Appendix A. It reads:

$$\langle e^{-2\beta w}\rangle \geq 1 + \frac{\left[\langle\,[\,f - \langle f\rangle\,]\,e^{-\beta w}\,\rangle\right]^2}{\langle\,[\,f - \langle f\rangle\,]^2\,\rangle} > 1, \quad (15)$$

where $f(x,p)$ is an arbitrary integrable function in the phase-space, and where

$$\langle f\rangle \equiv \int \mathrm{d}x\,\mathrm{d}p\,\mathcal{P}(x,p)\,f(x,p). \quad (16)$$

Eq. (15) is stronger than (14), since now $\langle e^{-2\beta w}\rangle$ is shown to be strictly larger than 1. Inequalities (14, 15) allow to understand how relevant the active realizations are with respect to both their probability and the amount of extracted work.

c. Dispersion of work. For sufficiently high temperatures one can make a cumulant expansion:

$$1 = \exp[-\beta\langle w\rangle + \frac{\beta^2}{2}(\langle w^2\rangle - \langle w\rangle^2) + ...] \quad (17)$$

which shows that for sufficiently high temperatures the ratio of the dispersion of work $\langle w^2\rangle - \langle w\rangle^2$ and its average increases with temperature:

$$\frac{\langle w^2\rangle - \langle w\rangle^2}{\langle w\rangle} = 2T. \quad (18)$$

A detailed survey of various cumulant expansion-based results derivable from BK equality is contained in Refs. [13, 14, 15].

F. Non-cyclic processes.

For non-cyclic processes there is an analog of equality (12), which is derived in similar way with the conclusion [21]: $\langle\, e^{-\beta w}\,\rangle = e^{-\beta(F(\tau)-F(0))}$, where $F = -T\ln Z$ is the corresponding free energy. This relation allows to calculate differences of free energy via (non-equilibrium) measurements of work. This generalized equality is not directly relevant for our present purposes, because here we are interested by the second law in Thomson's formulation which refers to cyclic processes.

III. QUANTUM ENSEMBLES AND THE DEFINITION OF FLUCTUATIONS OF WORK

A. The setup.

The quantum setup for studying thermally isolated processes is a straightforward extension of the classical one. (We denote all operators by a hat.)

An ensemble $\mathcal{E}$ of identically prepared quantum systems $\mathcal{S}$ is described at $t = 0$ by a density matrix $\hat{\rho}(0) = \hat{\rho}$. The eigenresolutions of $\hat{\rho}$ and of the Hamiltonian $\hat{H}$ read:

$$\hat{\rho} = \sum_{k=1}^{n} p_k |p_k\rangle\langle p_k|, \quad (19)$$

$$\hat{H} = \sum_{k=1}^{n} \varepsilon_k |\varepsilon_k\rangle\langle\varepsilon_k|, \quad (20)$$

where $\{|\varepsilon_k\rangle\}_{k=1}^n$ and $\{|p_k\rangle\}_{k=1}^n$ with $\langle\varepsilon_k|\varepsilon_l\rangle = \langle p_k|p_l\rangle = \delta_{kl}$ are the eigenvectors of $\hat{H}$ and $\hat{\rho}$, respectively, which form bases in the n-dimensional Hilbert space $\mathcal{H}$, and where ε_k and p_k are the corresponding eigenvalues.

Frequently, but not always, we will consider initially Gibbsian states:

$$\hat{\rho}(0) = \hat{\rho} = \frac{e^{-\beta\hat{H}}}{Z}, \quad Z = \mathrm{tr}\, e^{-\beta\hat{H}}, \quad (21)$$

$$p_k = \frac{e^{-\beta\varepsilon_k}}{\sum_{k=1}^n e^{-\beta\varepsilon_k}}, \quad |p_k\rangle = |\varepsilon_k\rangle, \quad k = 1,..,n, \quad (22)$$

where $T = 1/\beta \geq 0$ is the temperature of the ensemble. We shall order the eigenvalues of $\hat{H}$ as

$$\varepsilon_1 \leq \varepsilon_2 \leq ... \leq \varepsilon_n. \quad (23)$$

Then according to (22), the eigenvalues of $\hat{\rho}$ will be ordered as

$$p_1 \geq p_2 \geq ... \geq p_n > 0. \quad (24)$$

For the Gibbsian density matrix all eigenvalues are positive.

Analogously to the classical case, the Gibbsian state (21) is prepared for $t < 0$ by letting $\mathcal{S}$ to interact with a macroscopic thermal bath, and then decoupling it from

the bath, so that the interaction is absent for $t > 0$. There is, however, a relevant difference between quantum and classical: in the quantum situation the coupling of $\mathcal{S}$ with the bath has to be weak for the stationary state of $\mathcal{S}$ to be Gibbsian [4]. A detailed analysis of this and similar differences between the emergence of Gibbs distribution in quantum and classical situations is presented in [11, 12].

At $t = 0$ $\mathcal{S}$ starts to interact with an external macroscopic work source $\mathcal{W}$. The resulting evolution of $\mathcal{S}$ is generated by (an effective) Hamiltonian $\hat{H}\{R(t)\}$, which is time-dependent via classical (c-number) parameters $R(t)$. The evolution of $\mathcal{S}$ is thus unitary and has the same general features of reversibility as the dynamics of a completely isolated $\mathcal{S}$. It is well known that in general a Hamiltonian evolution of the complete system $\mathcal{S} + \mathcal{W}$ does not reduce to a Hamiltonian evolution for the state of $\mathcal{S}$. However, in the present case this is achieved owing to the *macroscopic* character of $\mathcal{W}$, as discussed in [4].

A cyclic process at the moment $t = \tau$ is defined in the same way as in classics, that is, via $R(\tau) = R(0)$, leading to

$$\hat{H}(\tau) = \hat{H}(0) = \hat{H}. \tag{25}$$

The Hamiltonian $\hat{H}(t)$ generates a unitary evolution:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = [\,\hat{H}(t), \hat{\rho}(t)\,], \tag{26}$$

$$\hat{\rho}(t) = \hat{U}_t\,\hat{\rho}(0)\,\hat{U}_t^\dagger, \tag{27}$$

$$\hat{U}_t = \overleftarrow{\exp}\left[-\frac{i}{\hbar}\int_0^t \mathrm{d}s\,\hat{H}(s)\right], \tag{28}$$

where $\overleftarrow{\exp}$ and $\overrightarrow{\exp}$ denote time-ordered and time-anti-ordered exponents, respectively.

B. Work.

The whole discussion in section II B directly applies in the quantum situation, except that $\mathcal{S}$ is now a quantum system, and Eqs. (4, 5) should be substituted by their quantum analogs (i.e., $\mathcal{P} \to \hat{\rho}$, $H \to \hat{H}$ and $\int \mathrm{d}x\,\mathrm{d}p \to \mathrm{tr}$). More specifically, the work W done by the external source $\mathcal{W}$ is identified with the average energy gained by the ensemble [3, 4]

$$W = \mathrm{tr}[\hat{\rho}(\tau)\,\hat{H} - \hat{\rho}\,\hat{H}] = \mathrm{tr}\,\hat{\rho}\,\hat{\Omega}, \tag{29}$$

where we denoted

$$\hat{\Omega} \equiv \hat{U}_\tau^\dagger\,\hat{H}(\tau)\,\hat{U}_\tau - \hat{H} = \hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau - \hat{H}. \tag{30}$$

Here $\hat{U}_\tau^\dagger\,\hat{H}(\tau)\,\hat{U}_\tau$ is the Hamiltonian operator in the Heisenberg representation at the end-time τ of the cyclic process. The operator $\hat{\Omega}$ is sometimes called 'operator of work' [6, 13, 27]. We shall show, however, in section IV A that it is not clear whether it fulfill all criteria to deserve this identification. Moreover, the much weaker interpretation of $\hat{\Omega}$ —by analogy to the classical expression (8)— as "energy difference operator in the Heisenberg representation" is also incorrect in general; see section IV A. In our approach $\hat{\Omega}$ will always appear inside averages over density matrices, so we do need any more particular interpretation of $\hat{\Omega}$; it will only enter the definition of work (29).

The remarks we made after Eq. (5) for the classical situation are valid in the quantum case as well. W is equal to the average energy decrease of the work source $\mathcal{W}$. This is a classical, mechanical energy which can transferred with 100% efficiency to other work-source, and, in particular, it can transferred to another mechanical degree of freedom performing classical deterministic motion. In that respect both the classical and quantum definitions are consistent and can be indistinguishable from the viewpoint of this mechanical degree. This property is the underlying reason why phenomenological thermodynamics, where not any (quantum or classical) identification of $\mathcal{S}$ is given, can exist.

The work is typically observed via suitable (classical) measurements done on the work source, or, alternatively, by measuring the initial and final average energies on the ensemble $\mathcal{E}$. Both these ways are routinely employed in practice, e.g., in NMR/ESR physics, where the system $\mathcal{S}$ corresponds to spin-$\frac{1}{2}$ under influence of external magnetic fields [50].

Finally, the quantum analog of formula (6) reads: $W = \int_0^\tau \mathrm{d}t\,\mathrm{tr}\left[\hat{\rho}(t)\,\frac{\mathrm{d}\hat{H}(t)}{\mathrm{d}t}\right]$, and Eq. (29) can be recovered from this formula upon integration by parts and using (26).

C. Quantum ensembles.

The definition of fluctuations of work in the classical situation was based on the distinction between classical ensemble of systems described by a probability distribution versus a single member of that ensemble. It should not be surprising that fluctuations of work in the quantum situation are closely tied to the meaning of what is a quantum ensemble.

Thus, for our further purposes we need an account of various features of quantum ensembles and their differences with respect to the classical ones. There are several sources in literature [43, 44, 45, 46, 47, 48] where this type of questions is studied with special attention [5].

[4] Due to weak coupling to the bath, the energy costs for switching the interaction on and off become negligible. This holds both in the quantum and the classical situation.

[5] Though the theory of quantum ensembles is almost as old as quantum mechanics itself, it still attracts lively discussions; see e.g. [64, 65, 66]. It is interesting to note the basic differences

1. Statistical interpretation of quantum mechanics.

Within the standard quantum mechanics a quantum 'state' is described by a density matrix $\hat{\rho}$. Any state, including a pure state $|\psi\rangle\langle\psi|$, describes an ensemble of identically prepared systems For instance, in an ideal Stern-Gerlach experiment all particles of the upper beam together are described by the wavefunction $|\uparrow\rangle$ or the pure density matix $|\uparrow\rangle\langle\uparrow|$. The description is optimal, in the sense that all particles have $\sigma_z = +1$, but incomplete in the sense that their σ_x and σ_y are unknown: upon measuring either of them, one will get ± 1 with equal probabilities.

2. Homogeneous ensembles.

In general, a density matrix $\hat{\rho}$ can be applied to describe two types of quantum ensembles, *homogeneous* and *inhomogeneous*.

For a homogeneous ensemble $\mathcal{E}(\hat{\rho})$ only the density matrix $\hat{\rho}$ is given and no further specification is made about a single system $\mathcal{S}$ from that ensemble. A typical example is an ensemble prepared by thermalization, that is, by letting each single system $\mathcal{S}$ to interact weakly with an equilibrium thermal bath, and waiting sufficiently long till the equilibrium state of $\mathcal{S}$ is established.

Let us study the defining feature of homogeneous ensembles in more details. We start by comparing them to classical ensembles. In the classical situation, the description of an ensemble by means of a probability distribution still implies that each single system has definite values for *all* its variables. For a homogeneous quantum ensemble $\mathcal{E}(\hat{\rho})$, only those observables (hermitian operators living in the Hilbert space $\mathcal{H}$) $\hat{A}$ that are dispersionless on $\mathcal{E}(\hat{\rho})$,

$$\left[\mathrm{tr}\left(\hat{A}\,\hat{\rho}\right)\right]^2 = \mathrm{tr}\left(\hat{A}^2\,\hat{\rho}\right), \tag{31}$$

can be said to have definite values for all single systems $\mathcal{S}$ from $\mathcal{E}(\hat{\rho})$. Indeed, it is shown in Appendix C that dispersionless observables satisfy

$$\hat{A}\,\hat{\rho} = \alpha\,\hat{\rho}, \tag{32}$$

where α is a c-number. This implies

$$\mathrm{tr}\,(\,\hat{A}^m\,\hat{\rho}\,) = \left[\mathrm{tr}\,\hat{A}\,\hat{\rho}\right]^m, \quad m = 0, 1, 2, 3..., \tag{33}$$

and the above statement follows. For a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$, we return from (32) to the standard notion of $|\psi\rangle$ being an eigenstate of $\hat{A}$.

Any other, non-dispersionless observable $\hat{B}$ — even if it commutes with the density matrix $\hat{\rho}$ — does not have a definite value in a single system $\mathcal{S}$ from $\mathcal{E}(\hat{\rho})$. It is true that for $[\hat{\rho}, \hat{B}] = 0$, $\mathcal{E}(\hat{\rho})$ can be prepared by mixing [6] pure states ensembles $\{\,\mathcal{E}(|p_k\rangle\langle p_k|)\,\}_{k=1}^n$ with probabilities $\{p_k\}_{k=1}^n$, where $\{\,|p_k\rangle\,\}_{k=1}^n$ and $\{p_k\}_{k=1}^n$ are, respectively, the common eigenvectors of $\hat{\rho}$ and $\hat{B}$ and the eigenvalues of $\hat{\rho}$. If $\mathcal{E}(\rho)$ is *known* to be prepared in such a way, then $\hat{B}$ has indeed definite values for each single member of $\mathcal{E}$. However, in general this need not apply, since there are (infinitely) many other ways to prepare the same ensemble $\mathcal{E}(\hat{\rho})$ via mixing N subensembles with density matrices $\{|\psi_\alpha\rangle\langle\psi_\alpha|\}_{\alpha=1}^N$ and probabilities $\{\lambda_\alpha\}_{\alpha=1}^N$. They correspond to the (infinitely) many ways in which the hermitian operator $\hat{\rho}$ can be decomposed as [45, 46, 47, 48]

$$\hat{\rho} = \sum_{\alpha=1}^{N} \lambda_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|, \quad \lambda_\alpha \geq 0, \quad \sum_{\alpha=1}^{N} \lambda_\alpha = 1, \tag{34}$$

where $|\psi_\alpha\rangle$ are some normalized — but in general not orthogonal— vectors living in the same n-dimensional Hilbert space $\mathcal{H}$ [7], and where $|\psi_\alpha\rangle\langle\psi_\alpha|$ are distinct.

The eigenresolution (19) is only a particular case of (34), and if now the ensemble $\mathcal{E}(\hat{\rho})$ was prepared by one of the ways corresponding to (34) with non-orthogonal $|\psi_\alpha\rangle$, the constituents of $\mathcal{E}(\hat{\rho})$ come from the subensembles $\mathcal{E}(|\psi_\alpha\rangle\langle\psi_\alpha|)\}$ and the observable $\hat{B}$ has in general no any definite value for these subensembles.

The above discussion allows to conclude with two related features of a homogeneous ensemble: *i)* a single member of such an ensemble does not by itself define a subensemble; *ii)* the ensemble cannot be thought to consist of definite subensembles.

3. Pure-state ensembles.

The description of a homogeneous ensemble via pure density matrices, $\hat{\rho}^2 = \hat{\rho}$, has several special features.

First of all, it is seen from (34) that for a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$ in the RHS of representation (34) only one term shows up: $|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|$ [8]. Thus, pure-state ensembles cannot be prepared via mixing of other ensembles of the

between classical and quantum ensembles were correctly understood by Elsasser as early as in 1937 [43].

[6] Mixing ensembles $\mathcal{E}(\hat{\rho}_1)$ and $\mathcal{E}(\hat{\rho}_2)$ with probabilities p_1 and p_2, respectively, means that one throws a dice with probabilities of outcomes equal to p_1 and p_2, and depending on the outcome one picks up a system from $\mathcal{E}(\hat{\rho}_1)$ or $\mathcal{E}(\hat{\rho}_2)$, keeping no information on where the system came from. Alternatively, one can join together Np_1 systems from $\mathcal{E}(\hat{\rho}_1)$ and Np_2 systems from $\mathcal{E}(\hat{\rho}_2)$ ($N \gg 1$), so that no information information is kept on where a single system came from. Then any subensemble of M systems ($N \gg M$) is described by the density matrix $\hat{\rho} = p_1\,\hat{\rho}_1 + p_2\,\hat{\rho}_2$. Note that the restriction $N \gg M$ is important, see, e.g., [65], and some confusion arose in literature for not taking it into account.

[7] Normalization and belonging to $\mathcal{H}$ are necessary for $|\psi_\alpha\rangle\langle\psi_\alpha|$ to describe some ensemble of the systems $\mathcal{S}$.

[8] This can also be deduced from a more general result: for any $|\psi_\alpha\rangle$ that can appear in (34), either $\hat{\rho} = |\psi_\alpha\rangle\langle\psi_\alpha|$, or $|\psi_\alpha\rangle$ is

system $\mathcal{S}$, or, put differently, pure-state ensembles are irreducible.

Second, this description is the maximally *complete* one possible in quantum mechanics. This known thesis can be substantiated as follows. First one notes from (31, 32) that for a fixed $\hat{\rho}$ dispersionless observables form a linear space: if two operators are dispersionless, so is their sum, and multiplication by a number conserves the dispersionless feature.

From (32) and Appendix C one sees that if the mixed density matrix $\hat{\rho}$ has k, $1 \leq k \leq n$, non-zero eigenvalues (n being the dimension of the Hilbert space $\mathcal{H}$), then the dimension of the linear space formed by the corresponding dispersionless observables is equal to

$$N_k = (n-k)^2 + 1. \tag{35}$$

This number is maximal for $k = 1$, that is, for pure density matrices. In other words, pure density matrices provide definite values for a larger set of observables than mixed density matrices [9]. For a mixed state all dispersionless observables have to be degenerate.

Though the above two features of irreducibility and completeness create a conceptual difference between pure and mixed density matrices, this should certainly not be taken as an invitation to prescribe pure density matrices to a single system, reserving the mixed ones for ensembles; further reasons for this are analyzed in Refs. [41, 44, 45, 46, 47, 48] [10].

4. Inhomogeneous ensembles.

A mixed density matrix $\hat{\rho}$ can also describe inhomogeneous ensembles. Such an ensemble $\mathcal{E}_{\rm i}$ is a collection of homogeneous subensembles $\{\, \mathcal{E}(\hat{\rho}_\alpha) \,\}_{\alpha=1}^N$ with probabilities $\{\, \lambda_\alpha \,\}_{\alpha=1}^N$, so that each single system from $\mathcal{E}_{\rm i}$ is known to be taken from the ensemble $\mathcal{E}(\hat{\rho}_\alpha)$ with probability λ_α, $\alpha = 1, .., N$. Obvious cases are when the subensembles $\mathcal{E}(\hat{\rho}_\alpha)$ are separated in space or in time, or by means of some other classical quantity.

Inhomogeneous ensembles are typically prepared by means of selective measurements [11]. In that case the above classical quantity is the corresponding record of the macroscopic apparatus by which this measurement was done. Below in section III E we describe in detail how an initially homogeneous ensemble can be separated into subensembles by means of a measurement.

The inhomogeneous ensemble $\mathcal{E}_{\rm i}$ is still described by the overall density matrix $\hat{\rho} = \sum_{\alpha=1}^N \lambda_\alpha \hat{\rho}_\alpha$, but in contrast to the homogeneous situation this is not the full description. The latter is provided by the list

$$\{\, \lambda_\alpha,\, \hat{\rho}_\alpha \,\}_{\alpha=1}^N. \tag{36}$$

So more information is known about the inhomogeneous ensemble $\mathcal{E}_{\rm i}$ then only $\hat{\rho}$. If the inhomogeneous ensemble is just a combination of homogeneous ones, this is obvious. If the inhomogeneous ensemble was prepared by means of a measurement, then the above information results from the measurement carried out and from selection of the outcomes (see more details in section III E below).

5. Prescribed ensemble fallacy.

This fallacy rests on forgetting the difference between homogeneous and inhomogeneous ensembles [46, 61]. In spite of explicit warnings [3], the fallacy frequently appears in applications and interpretations of quantum statistical physics. Consider, for example, the basic tool of statistical physics, the equilibrium ensemble described by the Gibbsian density matrix (21). It is typically obtained by thermalization process, that is, due to interaction with a thermal bath. One sometimes hears with respect to this ensemble that it represents the system being in states of definite energy with the corresponding probabilities p_k. This is a valid description of the ensemble only after the measurement of energy $\hat{H}$ has been done, something which is by itself not typical in applications. Moreover, as we recalled above and below, one can choose to make a different measurement, and then the interpretation in terms of definite energies will be explicitly wrong. The reason of why some applications —though starting from the above incorrect premise— do not lead to contradictions is clear: they use this premise merely for "explanations of what actually happens", while in real calculations and comparisons with experiment only the density matrix (21) is employed.

---

orthogonal to the linear space formed by the eigenvectors of $\hat{\rho}$ corresponding to eigenvalue zero. Indeed, let $|0\rangle$ be one such eigenvector, then $\langle 0|\hat{\rho}|0\rangle = \sum_\alpha \lambda_\alpha \, |\langle 0|\psi_\alpha\rangle|^2 = 0$; thus $|\langle 0|\psi_\alpha\rangle = 0$ for $\lambda_\alpha > 0$.

[9] For $k = n$ we get $N_k = 1$, since in this case only operators proportional to unity are dispersionless. For $n = 2$ and $k = 1$, $N_k = 2$: all dispersionless observables for a two-dimensional pure density matrix $|\psi\rangle\langle\psi|$ can be represented as $\alpha|\psi\rangle\langle\psi| + \beta|\psi_\perp\rangle\langle\psi_\perp|$, where $\langle\psi|\psi_\perp\rangle = 0$, and where α and β are two independent real numbers.

[10] Among reasons we find convincing is the analysis of the quantum measurement process [41].

[11] These measurements need not be done on the systems $\mathcal{S}$ directly, they can be indirect as well. Imagine an ensemble of two spin-$\frac{1}{2}$ particles described by pure density matrix $|\psi\rangle\langle\psi|$, where $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1 \otimes |+\rangle_2 \, + \, |-\rangle_1 \otimes |-\rangle_2)$, and where $|\pm\rangle_{1,2}$ are the eigenvectors of $\hat{\sigma}_z^{(1,2)}$ with eigenvalues ± 1 for the first and second particle, respectively. One can now measure $\hat{\sigma}_z^{(1)}$, and keep both the results of these measurements and the order of their appearance (thus, one keeps a sequence of random numbers ± 1). For the subensemble of the second spin this amounts to preparation of inhomogeneous ensemble $\{\frac{1}{2}, |+\rangle_2 \, {}_2\langle +| \; ; \; \frac{1}{2}, |-\rangle_2 \, {}_2\langle -|\}$.

D. Fluctuations of work.

Once the properties of quantum ensembles are clarified, we can proceed with the quantum definition of fluctuations of work. The most reasonable way to define this concept in the quantum situation is to proceed along the same lines as in classics, taking into account when needed the differences between quantum and classical ensembles.

It is convenient to separate the definition into the following steps.

1. The initial ensemble $\mathcal{E}(\hat{\rho})$ is homogeneous, since it was prepared by means of a thermal bath. With help of a suitable measurement (see section III E for details), one separates $\mathcal{E}(\hat{\rho})$ into irreducible, maximally complete subensembles $\{\,\mathcal{E}(|\psi_\alpha\rangle\langle\psi_\alpha|)\,\}_{\alpha=1}^{N}$ with probabilities $\{\,\lambda_\alpha\,\}_{\alpha=1}^{N}$, so that the resulting inhomogeneous ensemble is still described by the same density matrix $\hat{\rho}$ and thus (34) is valid.

In the quantum situation irreducible, maximally complete subensembles are described by pure density matrices $|\psi\rangle\langle\psi|$, as we recalled above. The important point is that these subensembles play here the same role as the single systems for the classical definition of fluctuations of work.

Note that once it is understood that the initial ensemble $\mathcal{E}(\hat{\rho})$ is homogeneous and that measurements are anyhow needed to make it inhomogeneous, we have to admit any measurement which will produce pure-state ensembles, even those with non-orthogonal $|\psi_\alpha\rangle$'s.

Recall that the present step of preparing an inhomogeneous ensemble out of the initial homogeneous one is absent in the classical situation, simply because there are no essentially inhomogeneous classical ensembles.

2. This step almost literally repeats its classical analog. The single systems from each subensemble $\mathcal{E}(|\psi_\alpha\rangle\langle\psi_\alpha|)$ interacts with the work source which realizes the same thermally isolated process on each single system from each subensemble.

The evolution of the corresponding subensemble during the cyclic process between times 0 and τ is given by the von Neumann equation

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_\alpha(t) = [\,\hat{H}(t), \hat{\rho}_\alpha(t)\,], \qquad \hat{\rho}_\alpha(0) = |\psi_\alpha\rangle\langle\psi_\alpha| \quad (37)$$

$$\hat{\rho}_\alpha(\tau) = \hat{U}_\tau\,\hat{\rho}_\alpha(0)\,\hat{U}_\tau^\dagger. \quad (38)$$

3. In analogy with the corresponding classical step we define the work w_α done on the subensemble α via Eq. (5):

$$\begin{aligned} w_\alpha \quad &= \mathrm{tr}\left(\hat{\Omega}\,|\psi_\alpha\rangle\langle\psi_\alpha|\right) \\ &= \langle\psi_\alpha(\tau)|\hat{H}|\psi_\alpha(\tau)\rangle - \langle\psi_\alpha(0)|\hat{H}|\psi_\alpha(0)\rangle. \end{aligned} \quad (39)$$

This is the average energy decrease of the mechanical degree of freedom of the work source due its interaction with the corresponding subensemble. Thus w_α has the meaning of work by itself, but it is a quantity that had to be averaged over the subensemble. The probability of w_α is equal to λ_α, since, as seen from (34), this is the probability by which the corresponding pure subensemble enters the overall ensemble described by $\hat{\rho}$.

Thus we defined a random c-number quantity work w with realizations w_α and probabilities λ_α:

$$w = \{w_\alpha\,,\,\lambda_\alpha\}_{\alpha=1}^{N}\,. \quad (40)$$

As follows from (29, 34) the work done on the overall ensemble is equal to the weighted average over the pure subensembles:

$$W = \sum_{\alpha=1}^{N}\lambda_\alpha w_\alpha. \quad (41)$$

Eq. (41) remains true for any initial ensemble. It is straightforward to see that our definition of fluctuations of work can be applied to any initial ensemble and not only to that described by the Gibbsian density matrix (21).

The thus defined fluctuations of work do depend on the pure ensembles $\{|\psi_\alpha\rangle\langle\psi_\alpha|\}_{\alpha=1}^{N}$, defined uniquely once the measurement separating the overall ensemble into pure subsensembles is specified. Strictly speaking, what we defined as fluctuations are the ones between subensembles (inter-subensemble fluctuations). Within the standard quantum theory we do not know how to define fluctuations of work inside of a irreducible subensemble. There were in literature some attempts in this direction, which are described in section VI. However, they do not satisfy the natural conditions on fluctuations of work, as outlined in the Introduction (arbitrary initial state; proper physical meaning). In particular, the approach based on the "operator of work" is not applicable, since we will explain that this operator does not satisfy the proper criteria [12].

If there is no interaction with any work source, that is, the Hamiltonian $\hat{H}$ is time-independent, and if in addition $[\hat{\rho}, \hat{H}] = 0$, then the ensemble described by $\hat{\rho}$ is stationary: all (one-time) averages are time-independent. Now note that the stationary ensemble can be decomposed into non-stationary subensembles, since in general $[\,|\psi_\alpha\rangle\langle\psi_\alpha|\,,\,\hat{H}] \neq 0$. This is clearly impossible for a classical ensemble, but in the context of fluctuations of work this fact implies nothing pathologic, since work is defined for any initial ensemble, not only for stationary ones. It is checked from (39) that if there is no interaction with the work-sources, then $\Omega \equiv 0$, and all possible realizations of work are zero.

4. Note that for macroscopic systems it is not realistic to have available measurements producing pure-

[12] Thus if these fluctuations exist, and we assume they do, their description seems to be outside of today's theories. It might be of some interest to see whether more detailed definitions of fluctuations of work can be given in theories of subquantum mechanics, e.g., Bohmian or Nelsonian mechanics.

state subensembles, since the directly availabale measurements are only those of macroscopic quantities which are typically degenerate. In this case we may need to apply a coarse-grained definition of fluctuations of work, where the initial mixed ensemble is separated into mixed subensembles described by density matrices $\hat{\sigma}_\gamma$ ($\hat{\sigma}_\gamma^2 \neq \hat{\sigma}_\gamma$)

$$\hat{\rho} = \sum_\gamma \nu_\gamma \hat{\sigma}_\gamma, \quad \nu_\gamma \geq 0, \quad \sum_\gamma \nu_\gamma = 1. \tag{42}$$

The definition then proceeds as above with obvious changes (e.g., $|\psi_\alpha\rangle\langle\psi_\alpha| \to \hat{\sigma}_\gamma$ in (39)).

This is a coarse-grained definition, since the realizations of work $\mathrm{tr}\,(\hat{\Omega}\,\hat{\sigma}_\gamma)$ can be reduced to more fundamental ones, i.e., each of them can be presented as a convex sum of $\mathrm{tr}\,(|\psi_\alpha\rangle\langle\psi_\alpha|\,\hat{\Omega}\,)$. As a consequence fluctuations of work —as quantified, e.g., by dispersion of work defined and discussed in section IV— are maximal for pure-state decompositions (more details on this are found in section IV A).

E. Separation of a homogeneous ensemble into pure subensembles by filtering outcomes of a POVM measurement.

1. *Positive Operator Valued Measurements.*

It is now our purpose to discuss how precisely one separates with help of measurements an initial homogeneous ensemble $\mathcal{E}(\hat{\rho})$ into pure (necessarily homogeneous) subensembles.

The most general type of a quantum measurement corresponds to Positive Operator Valued Measure (POVM) [46, 47] defined via N operators $\hat{G}_\alpha$ — not necessarily orthogonal — living in the n-dimensional Hilbert space $\mathcal{H}$ and satisfying the completeness relation

$$\sum_{\alpha=1}^{N} \hat{G}_\alpha^\dagger \hat{G}_\alpha = \hat{1}. \tag{43}$$

The most standard measurements of an observable $\hat{A}$ living in the n-dimensional Hilbert space $\mathcal{H}$ and having non-degenerate spectrum $\{a_\alpha\}_{\alpha=1}^n$ are included in (43), since now $N = n$ and $\{\hat{G}_\alpha\}_{\alpha=1}^n = \{|a_\alpha\rangle\langle a_\alpha|\}_{\alpha=1}^n$, where the latter is the set of orthonormal eigenvectors of $\hat{A}$. If the spectrum of $\hat{A}$ happens to have degeneracies, so that each eigenvalue a_α has multiplicity n_α, then $\hat{G}_\alpha$ is the n_α-dimensional projector on the subspace formed by n_α linearly independent eigenvectors of $\hat{A}$ which correspond to the eigenvalue a_α. Here $N \leq n$ is equal to the number of distinct eigenvalues of $\hat{A}$.

If the measurement described by (43) is done on the ensemble described by a density matrix $\hat{\rho}$, then the result α is found with probability

$$\lambda_\alpha = \mathrm{tr}(\hat{G}_\alpha^\dagger \hat{G}_\alpha \hat{\rho}) = \mathrm{tr}(\hat{G}_\alpha \hat{\rho} \hat{G}_\alpha^\dagger), \tag{44}$$

where $\lambda_\alpha \geq 0$ and $\sum_{\alpha=1}^N \lambda_\alpha = 1$, due to (43). After selecting results of the measurements referring to the outcome α one has the (sub)ensemble of systems described by a density matrix

$$\hat{\rho}'_\alpha = \frac{\hat{G}_\alpha \hat{\rho} \hat{G}_\alpha^\dagger}{\mathrm{tr}(\hat{G}_\alpha^\dagger \hat{G}_\alpha \hat{\rho})}. \tag{45}$$

This subensemble occura with probability λ_α as given by (44), simply because this is the probability of the outcome α. The overall post-measurement inhomogeneous ensemble thus *consists of* N subensembles each of which has a density matrix (45) and probability (44). The density matrix of the overall post-measurement ensemble is

$$\hat{\rho}' = \sum_{\alpha=1}^{N} \lambda_\alpha \hat{\rho}'_\alpha. \tag{46}$$

POVM's are related to more usual projective measurements, where $\hat{G}_\alpha$ are mutually orthogonal, $\hat{G}_\alpha \hat{G}_\beta = \hat{G}_\alpha \delta_{\alpha\beta}$, projections into eigenspaces of some hermitian operator: according to Neumark's theorem [46, 47] every POVM can be realized as some (non-unique) projective measurement in a larger Hilbert space, that is involving additional degrees of freedom. A detailed discussion of this theorem and various versions is given in [46, 47]. In Appendix D we shall discuss an example of it that is relevant for our purposes.

2.

Applying a POVM measurement, one now wishes to separate the mixed quantum ensemble described by the density matrix $\hat{\rho}$ into pure subensembles. The density matrix $\hat{\rho}'$ of the overall post-measurement ensembles should then coincide with $\hat{\rho}$ given in (19) or (21), while $\hat{\rho}'_\alpha$ appearing in (45) should be pure:

$$\hat{\rho}'_\alpha = |\psi_\alpha\rangle\langle\psi_\alpha|. \tag{47}$$

Then the density matrix (21) is decomposed as in (34).

Let us first see which $\{\lambda_\alpha\}_{\alpha=1}^N$ and $\{|\psi_\alpha\rangle\langle\psi_\alpha|\}_{\alpha=1}^N$ are allowed to enter in (34), and then we shall discuss which specific measurements should be done to achieve the actual separation.

It will prove useful to write (34) in an equivalent way

$$\hat{\rho} = \sum_{\alpha=1}^{N} |\widetilde{\psi_\alpha}\rangle\langle\widetilde{\psi_\alpha}|, \quad |\widetilde{\psi_\alpha}\rangle \equiv \sqrt{\lambda_\alpha}|\psi_\alpha\rangle, \tag{48}$$

since it will allow us to focus on $|\widetilde{\psi_\alpha}\rangle$, keeping in mind that the probabilities λ_α can always be recovered via $\lambda_\alpha = \langle\widetilde{\psi_\alpha}|\widetilde{\psi_\alpha}\rangle$.

According to the ensemble classification theorem [51, 52, 53, 54], if one has

$$|\widetilde{\psi_\alpha}\rangle = \sum_{k=1}^{n} M_{\alpha k} \sqrt{p_k}|p_k\rangle, \tag{49}$$

where $\{p_k\}_{k=1}^n$, $\{|p_k\rangle\}_{k=1}^n$ are the eigenvalues and eigenfunctions of the density matrix $\hat{\rho}$, and where $M_{\alpha k}$ are complex numbers satisfying

$$\sum_{\alpha=1}^{N} M_{\alpha k} M_{\alpha j}^{*} = \delta_{ij}, \quad k,j = 1,...,n, \tag{50}$$

then Eq. (48) becomes $\hat{\rho} = \sum_{k=1}^n p_k |p_k\rangle\langle p_k|$, as it should [13].

The converse appears to be true as well: any decomposition (48) admits a representation (49) with some complex numbers $M_{\alpha k}$ satisfying (50) [14].

Note that Eq. (49) implies the following formula for the probabilities λ_α:

$$\lambda_\alpha = \widetilde{\langle\psi_\alpha|}\widetilde{|\psi_\alpha\rangle} = \sum_{k=1}^{n} |M_{\alpha k}|^2 p_k. \tag{51}$$

As seen from (50), the very possibility of writing Eq. (48) implies

$$N \geq n, \tag{52}$$

since $M_{\alpha k}$ can be viewed as n different N-component orthonormal vectors. The rectangular matrix $\{\{M_{\alpha k}\}_{\alpha=1}^N\}_{k=1}^n$ can be completed to a unitary $N \times N$ matrix by adding suitable elements.

It is now straightforward to see which POVM can be taken to achieve the decomposition (48). Take, for example,

$$G_\alpha = \frac{\widetilde{|\psi_\alpha\rangle}\widetilde{\langle\psi_\alpha|}\,\hat{\rho}^{-1/2}}{\sqrt{\widetilde{\langle\psi_\alpha}\,|\,\widetilde{\psi_\alpha\rangle}}} = \sqrt{\lambda_\alpha}\,|\psi_\alpha\rangle\,\langle\psi_\alpha|\,\hat{\rho}^{-1/2}, \tag{53}$$

where $|\psi_\alpha\rangle$ is defined in (48). Note that the converse appears to be true as well. For given POVM (43) with

$$\hat{G}_\alpha^\dagger \hat{G}_\alpha = |\pi_\alpha\rangle\langle\pi_\alpha|, \tag{54}$$

where $|\pi_\alpha\rangle$ satisfying

$$\hat{1} = \sum_{\alpha=1}^{N} |\pi_\alpha\rangle\langle\pi_\alpha|, \tag{55}$$

have to be neither orthogonal, nor normalized [15], one can construct a representation (48, 34) of $\hat{\rho}$ as

$$\hat{\rho} = \sum_{\alpha=1}^{N} \hat{\rho}^{1/2}\,|\pi_\alpha\rangle\langle\pi_\alpha|\,\hat{\rho}^{1/2}. \tag{56}$$

Thus, we have seen how all possible decompositions of a mixed ensemble into pure subensenbles can be constructed via suitable measurements.

We stress that the decompositions into a specific set of subensembles is related to a physical measurement, rather than to a mathematical choice.

3. Preparation versus measurements.

To avoid possible confusions we recall once again that the above separation procedure corresponds to *preparation* of the inhomogeneous ensemble $\{\lambda_\alpha, \mathcal{E}(|\psi_\alpha\rangle\langle\psi_\alpha|)\}_{\alpha=1}^N$ with $\hat{\rho} = \sum_\alpha \lambda_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$, starting from the initial homogeneous ensemble $\mathcal{E}(\hat{\rho})$. Though this preparation was based on a suitable measurement process, we were not interested by some aspects usually associated with it. For example, we did not keep track of the pointer variable of the measuring apparatus, which obviously should be the main goal of any measurement process studied for its own purposes [41]. We were more interested by the influence of the measurement process on the final state of the system $\mathcal{S}$, which is the basic characteristic feature of the preparation process in quantum mechanics [47].

F. Discussion.

There are several questions on the physical meaning of the proposed definition of fluctuations of work that we decided to discuss separately.

Question 1. Among all decompositions (34) of the Gibbsian density matrix $\hat{\rho}$, there is a unique one (up to accidental degeneracies of the spectrum) given by the eigenvectors of $\hat{\rho}$ and realized via measurement of the Hamiltonian $\hat{H}$. Then the energy has a definite value on each subensemble. Should not one therefore restrict the definition of fluctuations of work to this separation only?

Answer 1. There are at least two reasons why the answer is no. First, even if the energy has a definite value initially, it will in general not have any definite value

---

[13] Note that any vector $\widetilde{|\psi_\alpha\rangle}$ having $\widetilde{\langle\psi_\alpha|}\widetilde{|\psi_\alpha\rangle} < 1$ and living in the Hilbert space formed by the eigenvectors of $\hat{\rho}$ corresponding to its non-zero eigenvalues, can appear in at least one separation (48) of $\hat{\rho}$. This follows from (49).

[14] To prove this part of the statement, recall Footnote 8, expand $\widetilde{|\psi_\alpha\rangle}$ over the eigenbase $|p_k\rangle$ of $\hat{\rho}$: $\widetilde{|\psi_\alpha\rangle} = \sum_{k=1}^n \langle p_k|\widetilde{\psi_\alpha}\rangle\,|p_k\rangle$, substitute this into (48), and then deduce (50) using the orthonormality and completness of the above base in the Hilbert space $\mathcal{H}$: $\sum_{\alpha=1}^N \langle p_k|\widetilde{\psi_\alpha}\rangle\,\widetilde{\langle\psi_\alpha}|p_l\rangle = \langle p_k|\hat{\rho}|p_l\rangle = \delta_{kl}\,p_k$. Thus, any decomposition (48, 34) can be constructed via (49) and $M_{\alpha k}\langle p_k|\widetilde{\psi_\alpha}\rangle/\sqrt{p_k}$ satisfying (49). If some eigenvalues of $\hat{\rho}$ are equal to zero, than the above construction should be restricted to eigenvectors of $\hat{\rho}$ corresponding to its non-zero eigenvalues.

[15] If one assumes in Eq. (55) that $|\pi_\alpha\rangle$ are normalized, $\langle\pi_\alpha|\pi_\alpha\rangle = 1$, then this leads to orthogonality: $\langle\pi_\beta|\pi_\alpha\rangle = \delta_{\alpha\beta}$. Indeed, denoting $\hat{\Pi}_\alpha = |\pi_\alpha\rangle\langle\pi_\alpha|$, one gets $\sum_{\alpha\neq\beta}^N \left(\hat{\Pi}_\alpha\hat{\Pi}_\beta\right)^\dagger \left(\hat{\Pi}_\alpha\hat{\Pi}_\beta\right) = \sum_{\alpha\neq\beta}^N \hat{\Pi}_\beta\hat{\Pi}_\alpha\hat{\Pi}_\alpha\hat{\Pi}_\beta = \hat{\Pi}_\beta(1 - \hat{\Pi}_\beta)\hat{\Pi}_\beta = 0$. Since $\left(\hat{\Pi}_\alpha\hat{\Pi}_\beta\right)^\dagger \left(\hat{\Pi}_\alpha\hat{\Pi}_\beta\right)$ is non-negative by construction, one concludes $\hat{\Pi}_\alpha\hat{\Pi}_\beta = 0$ for $\alpha \neq \beta$.

at the final moment, since an eigenvector of the initial Hamiltonian may evolve into a superposition of eigenvectors. Thus, there are no special reasons to insist on the feature of separating with respect to energy. Second, more general separations are anyhow necessary to define fluctuations of work for an arbitrary ensemble, which cannot be decomposed into subensembles with each of them having a definite value of energy.

Question 2. Is the orthogonal separation not special by the fact that various ensembles are described by orthogonal pure density matrices, and can thus be discriminated unambigiously?

Answer 2. By definition any POVM is connected with an unambigious descrimination of its different outcomes. This can be additionally clarified by looking at the example of the projective realization of a POVM presented in Appendix D, where various subensembles constructed after the measurement are seen to be described by orthogonal wave-functions in the composite Hilbert space $\mathcal{H} \otimes \mathcal{H}'$. The above question mixes the present situation with a different one, where one is given a single system coming from one of two ensembles having non-orthogonal density matrices, and is requested to determine by means of a measurement from which ensemble it is coming. Then, indeed, no measurements can ensure unambigious discrimination [46].

Question 3. The authors prescribe to the viewpoint that even pure density matrices (wave functions) describe an ensemble of quantum systems and not a single system, as some people like to think. How the proposed definition will change, if one would wish to insist on the latter interpretation of quantum mechanics?

Answer 3. The necessity of prescribing even the pure density matrices to ensembles of quantum systems was stressed in [44, 45, 46, 47]. In particular, it is needed for the consistent solution of the quantum measurement problem [41, 47]. It is also known that with respect to certain aspects of quantum theory the prescription of pure density matrices to a single system is relatively harmless. We do not have space to discuss in detail what are those aspects and what precisely is meant by "relatively harmless". At any rate, we do not advise to make the latter prescription, and the readers who wish to do that have to proceed on their own risk. We may mention that the definition of fluctuations of work remains then basically unchanged, but becomes conceptually closer to its classical analog, since now in defining fluctuations of work one assumes to operate with single systems both in quantum and classical situations.

IV. DISPERSION OF WORK.

The most direct quantity that characterizes how the realizations (39) of the random quantity work are spread around their mean $W = \sum_{\alpha=1}^{N} \lambda_\alpha w_\alpha$, is the (inter-subensemble) dispersion

$$\delta w^2 = \sum_{\alpha=1}^{N} \lambda_\alpha \left[\langle \psi_\alpha | \hat{\Omega} | \psi_\alpha \rangle - \mathrm{tr}(\hat{\Omega}\hat{\rho}) \right]^2 \tag{57}$$

$$= \sum_{\alpha=1}^{N} \lambda_\alpha (w_\alpha - W)^2 = \sum_{\alpha=1}^{N} \lambda_\alpha w_\alpha^2 - W^2. \tag{58}$$

In contrast to W, this quantity depends explicitly on the subensembles used to define w_α in (39). So it depends explicitly on the physical process that separated the initial ensemble into subensembles.

It is useful to to determine the maximal $\delta w^2_{\max}$ and the minimal $\delta w^2_{\min}$ values of δw^2 over all possible decompositions $\{|\psi_\alpha\rangle\langle\psi_\alpha|, \lambda_\alpha\}_{\alpha=1}^N$ corresponding to the fixed $\hat{\rho} = \sum_{\alpha=1}^N \lambda_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$. According to (56) these extremizations can be equally well carried out over all possible decompositions of unity in our n-dimensional Hilbert space,

$$\sum_{\alpha=1}^{N} \hat{\Pi}_\alpha = \hat{1}, \qquad \hat{\Pi}_\alpha = |\pi_\alpha\rangle\langle\pi_\alpha|, \tag{59}$$

where $\{|\pi_\alpha\rangle\}_{\alpha=1}^N$ have in general to be neither normalized nor orthogonal.

Note that dispersions similar to (57), with Ω corresponding to some other relevant observable, where introduced and studied in quantum optics, where separation of an ensemble by means of (continuous) measurements are well-known and were studied both experimentally and theoretically; see [48] for a review. The results we present below on the minimal and maximal values of the dispersion δw^2 do not depend on the details of Ω and can thus be useful in general.

A. Maximal dispersion of work.

The maximization of δw^2 over all possible separations (34, 59) for given $\hat{\rho}$ and $\hat{\Omega}$ is carried out in Appendix G. The result is

$$\delta w^2_{\max} = \sum_{i,k=1}^{n} \frac{2p_i p_k}{p_i + p_k} |\langle \varepsilon_k | \hat{\Omega} | \varepsilon_i \rangle|^2 - W^2 \tag{60}$$

$$= 2 \int_0^\infty \mathrm{d}s \, \mathrm{tr} \left[\left(\hat{\Omega}\, \hat{\rho}\, e^{-s\hat{\rho}} \right)^2 \right] - W^2. \tag{61}$$

This maximum is reached on $\{|\pi_\alpha\rangle\}_{\alpha=1}^n$ being the eigenvectors of a hermitian operator

$$\hat{X} = \sum_{i,k=1}^{n} \frac{2p_i p_k \ \langle p_i | \hat{\Omega} | p_k \rangle}{p_i + p_k} |p_i\rangle\langle p_k|, \tag{62}$$

where p_k and $|p_k\rangle$ are the eigenvalues and eigenvectors of $\hat{\rho}$, as defined by (19).

Only when $\hat{\rho}$ and $\hat{\Omega}$ commute, $[\hat{\rho}, \hat{\Omega}] = 0$, the maximal dispersion (60) reduces to the more usual expression

$\mathrm{tr}[\hat\rho\,\hat\Omega^2] - [\mathrm{tr}(\hat\rho\,\hat\Omega)]^2$. This and related questions are discussed in more detail around Eqs. (101, 102).

The maximal dispersion (60, 61) provides an upper bound for the dispersion of work defined in a coarse-grained way; see the discussion around Eq. (42). Indeed according to that discussion the coarse-grained dispersion of work defined with respect to separation of $\mathcal{E}(\hat\rho)$ to mixed-state subensembles reads

$$\delta w^2_{\rm cg} = \sum_\gamma \nu_\gamma (\mathrm{tr}(\hat\sigma_\gamma\hat\Omega) - W)^2. \tag{63}$$

Note a decomposition of $\hat\sigma_\gamma$ into some set of pure-state subensembles, $\hat\sigma_\gamma = \sum_\alpha \mu^{(\gamma)}_\alpha |\psi^{(\gamma)}_\alpha\rangle\langle\psi^{(\gamma)}_\alpha|$, where $\mu^{(\gamma)}_\alpha$ are the corresponding probabilities with $\sum_\alpha \mu^{(\gamma)}_\alpha = 1$. One now finds that the dispersion δw^2 defined as in Eqs. (57, 58), that is, via the separation of the ensemble $\mathcal{E}(\hat\rho)$ into pure-state subensembles $\hat\rho = \sum_{\alpha,\gamma}\nu_\gamma\mu^{(\gamma)}_\alpha|\psi^{(\gamma)}_\alpha\rangle\langle\psi^{(\gamma)}_\alpha|$, is always not smaller than $\delta w^2_{\rm cg}$:

$$\begin{aligned}\delta w^2 - \delta w^2_{\rm cg} \;&\geq\; \sum_{\alpha,\gamma}\nu_\gamma\mu^{(\gamma)}_\alpha\,(\,\langle\psi^{(\gamma)}_\alpha|\hat\Omega|\psi^{(\gamma)}_\alpha\rangle - W\,)^2 - \delta w^2_{\rm cg}\\ &= \sum_{\alpha,\gamma}\nu_\gamma\mu^{(\gamma)}_\alpha\,\Big(\,\langle\psi^{(\gamma)}_\alpha|\hat\Omega|\psi^{(\gamma)}_\alpha\rangle\\ &\quad - \sum_\beta \mu^{(\gamma)}_\beta\langle\psi^{(\gamma)}_\beta|\hat\Omega|\psi^{(\gamma)}_\beta\rangle\Big)^2 \geq 0.\end{aligned} \tag{64}$$

1. *The behavior of the maximal dispersion $\delta w^2_{\max}$ for high and low temperatures.*

With $\hat\rho$ given by the Gibbs distribution (21, 19, 22), one gets from Eq. (60)

$$\delta w^2_{\max} \to 0, \quad \text{for} \quad T \to 0, \tag{65}$$

where T is the temperature of the Gibbsian ensemble. This is a natural result, as for a finite system $\mathcal{S}$ and $T \to 0$ one gets $\hat\rho \to |\varepsilon_0\rangle\langle\varepsilon_0|$, where according to (20, 19), $|\varepsilon_0\rangle$ is the common eigenvector of $\hat\rho$ and $\hat H$ corresponding to the lowest energy (assuming that the latter is not degenerate). As no separation of a pure state into subensembles is possible, the work can take only one value. It is obvious that this is a general feature: the work does not fluctuate if the initial ensemble is pure. In the same way as in classics, fluctuations of work are present for mixed ensembles only. In this respect the dispersion of work is similar to the von Neumann entropy $S_{\rm vN} = -\mathrm{tr}\hat\rho\ln\hat\rho$, which is also equal to zero for pure density matrices $\hat\rho$.

For very high temperatures, where $\hat\rho \simeq \hat 1/n$, one gets from (60)

$$\delta w^2_{\max} = \frac{1}{n}\,\mathrm{tr}\left(\hat\Omega^2\right). \tag{66}$$

It is seen that for high temperatures the maximal dispersion may be $\mathcal{O}(1)$, provided that the (positive) eigenvalues of $\hat\Omega^2$ are finite and do not scale with n.

B. Minimal dispersion of work.

Here we show that there are decompositions into subensembles such that for any $\alpha = 1, ..., N$:

$$w_\alpha = \langle\psi_\alpha|\hat\Omega|\psi_\alpha\rangle = \sum_{\beta=1}^N \lambda_\beta w_\beta = W, \tag{67}$$

that is, the work does not fluctuate at all. In particular, this means that the dispersion δw^2 attains its minimal value equal to zero. This fact is contrasting to the classical situation, where according to points b and c in section II E, $w(x,p)$ should be negative at least for some values of (x,p), and the dispersion of work is large at least for sufficiently high temperatures.

Recall that due to the parametrization (54, 56, 59), Eq. (67) can be written as

$$\frac{\langle\pi_\alpha|\hat\rho^{1/2}\,\hat\Omega\,\hat\rho^{1/2}|\pi_\alpha\rangle}{\langle\pi_\alpha|\hat\rho|\pi_\alpha\rangle} = \mathrm{tr}(\,\hat\rho\,\hat\Omega\,), \tag{68}$$

where $\{|\pi_\alpha\rangle\}_{\alpha=1}^N$ with $N \geq n$ have to satisfy (59). This is equivalent to

$$0 \;=\; \langle\pi_\alpha|\hat Y|\pi_\alpha\rangle, \tag{69}$$

$$\hat Y \;\equiv\; \hat\rho^{1/2}\,\hat\Pi_\alpha\,\hat\rho^{1/2} - \mathrm{tr}\,[\,\hat\Omega\,\hat\rho\,]\,\hat\rho, \tag{70}$$

where $\hat Y$ is hermitian and traceless:

$$\mathrm{tr}\,\hat Y = 0. \tag{71}$$

We now intend to show that in the Hilbert space $\mathcal{H}$ there are orthonormal bases $\{|\pi_i\rangle\}_{i=1}^n$ which for the given $\hat Y$ do satisfy to (67, 69).

1. *Some concepts from majorization theory.*

To this end, let us recall some concepts from the mathematical theory of majorization [56, 57, 58, 59]. For two real vectors $x = (x_1 \geq ... \geq x_n)$ and $y = (y_1 \geq ... \geq y_n)$, with their components arranged in non-increasing way, y is said to majorize x,

$$x \prec y, \tag{72}$$

if the following conditions are satisfied

$$\sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i, \qquad k = 1, .., n-1, \tag{73}$$

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i. \tag{74}$$

Due to Horn's theorem [56, 57, 58, 59], Eq. (72) implies the existence of a $n \times n$ unitary matrix Q_{ij} such that

$$x_i = \sum_{k=1}^n y_j\,|Q_{ij}|^2. \tag{75}$$

The proof of this statement is recalled in Appendix F. This proof is constructive, since it allows to gets Q_{ij} starting from given x and y.

2. The minimal dispersion of work is zero.

Now denote by $(y_1 \geq ... \geq y_n)$ the eigenvalues of the hermitian matrix $\hat{Y}$ arranged in non-increasing way. Denote by $\{|y_i\rangle\}_{i=1}^n$ the corresponding eigenvectors. As follows from (71, 73, 74)

$$(y_1, ..., y_n) \succ (0, ..., 0). \tag{76}$$

According to (75) there exists a unitary operator $\hat{Q}$ in the Hilbert space $\mathcal{H}$ such that

$$0 = \sum_{j=1}^{n} y_j \, |\langle y_j|\hat{Q}|y_i\rangle|^2 = \langle y_i| \, \hat{Q}^\dagger \, \hat{Y} \, \hat{Q} \, |y_i\rangle. \tag{77}$$

By denoting

$$\hat{Q} \, |y_i\rangle = |\pi_i\rangle, \qquad i = 1, ..., n, \tag{78}$$

we see that (77) and the desired statement (69) are equivalent.

C. Dispersion of work averaged over all separations of the ensemble.

We have obtained the maximal and the minimal values of the dispersion of work δw^2. It is useful to have a third characteristic value of δw^2, the dispersion of work for a randomly chosen separation of the initial ensemble described by $\hat{\rho}$ into pure subensembles. Such a quantity will not depend explicitly on the measurement used for separation, and thus will help to understand how typical are the maximal and the minimal values of δw^2.

Note from Eqs. (49, 50) that for a given separation of $\hat{\rho}$, that is, for a given representation (34), the pure density matrices $|\psi_\alpha\rangle\langle\psi_\alpha|$ are expressed via elements $M_{\alpha i}$ of a $N \times N$ unitary matrix M (see the remark after (52)). We shall define the average dispersion $\delta w^2_{\rm av}$ by assuming that M is random, and then integrating $\delta w^2\{M_{\alpha i}\}$ over all possible unitary $N \times N$ matrices. Since there are no reasons for introducing *a priori* biases, we shall assume for the above integration the most uniform, unitary-invariant measure (Haar's measure):

$$\delta w^2_{\rm av} = \frac{\int \prod_{i,\alpha=1}^{N} {\rm d}\Re M_{\alpha i} \, {\rm d}\Im M_{\alpha i} \, \Theta\{M_{\alpha i}\} \, \delta w^2\{M_{\alpha i}\}}{\int \prod_{i,\alpha=1}^{N} {\rm d}\Re M_{\alpha i} \, {\rm d}\Im M_{\alpha i} \, \Theta\{M_{\alpha i}\}}, \tag{79}$$

where $\Theta\{M_{\alpha i}\}$ comes due to the unitarity constraint

$$\Theta\{M_{\alpha i}\} = \prod_{\alpha=1}^{N} \delta \left[\sum_{i=1}^{N} |M_{\alpha i}|^2 - 1 \right] \prod_{\alpha<\beta}^{N} \delta \left[\sum_{i=1}^{N} M_{\alpha i} M^*_{\beta i} \right].$$

The rows (or, equivalently, the columns) of the matrix M are thus assumed to be a set of N orthonormalized, uniformly random vectors. The quantity $\delta w^2_{\rm av}$ is calculated in Appendix H:

$$\delta w^2_{\rm av} = \int_0^\infty {\rm d}s \left[\prod_{k=1}^{n} \frac{1}{1+sp_m} \right] \left[\sum_{i=1}^{n} \left(\frac{p_i \, \langle \varepsilon_i|\hat{\Omega}|\varepsilon_i\rangle}{1+sp_i} \right)^2 + \left(\sum_{i=1}^{n} \frac{p_i \, \langle \varepsilon_i|\hat{\Omega}|\varepsilon_i\rangle}{1+sp_i} \right)^2 \right] - W^2. \tag{80}$$

Note that $\delta w^2_{\rm av}$ depends neither on N, nor on the off-diagonal elements $\langle \varepsilon_i|\hat{\Omega}|\varepsilon_j\rangle$ of $\hat{\Omega}$.

For $\hat{\rho}$ having the Gibbsian form (21, 19, 22), $\delta w^2_{\rm av}$ has the following features for low and high temperatures T. It goes to zero for $T \to 0$ for the same reasons as $\delta w^2_{\rm max}$ does. In contrast, for very high temperatures, where $\hat{\rho} \simeq 1/n$, one has from (80)

$$\delta w^2_{\rm av} = \frac{1}{n(n+1)} \sum_{i=1}^{n} \langle \varepsilon_i|\hat{\Omega}|\varepsilon_i\rangle^2. \tag{81}$$

Under the same natural condition that we adopted for studying the high-temperature behavior of $\delta w^2_{\rm max}$, that is, $\langle \varepsilon_i|\hat{\Omega}|\varepsilon_i\rangle$ are finite and do not scale with n, we see that $\delta w^2_{\rm av} \propto 1/n$ for $n \gg 1$, which is a typical behavior for dispersions of fluctuating macroscopic quantities in statistical physics [3]. Note the difference with the high-temperature behavior of the maximal dispersion given by Eq. (66).

D. The maximal and the average dispersion of work illustrated for a two-level system.

Let us give concrete expressions of $\delta w^2_{\rm max}$ and $\delta w^2_{\rm av}$ for a two-level system $\mathcal{S}$. The initial Gibbsian density matrix is now a 2×2 diagonal matrix with eigenvalues p_1 and $p_2 \leq p_1$ as given by (19). The most general matrix form of the traceless and hermitian operator $\hat{\Omega}$ in this two-dimensional situation is

$$\hat{\Omega} = \begin{pmatrix} \omega & \chi \\ \chi^* & -\omega \end{pmatrix}. \tag{82}$$

Eqs. (60, 80) produce then the following expressions for $\delta w^2_{\rm max}$ and $\delta w^2_{\rm av}$, respectively:

$$\delta w^2_{\rm max} = \omega^2(1-x^2) \left(1 + \frac{|\chi|^2}{\omega^2} \right), \tag{83}$$

$$\delta w^2_{\rm av} = \omega^2(1-x^2) \left[1 - \frac{1}{x^2} \left(1 + \frac{1-x^2}{2x} \, \ln \frac{1-x}{1+x} \right) \right], \tag{84}$$

where

$$x \equiv p_1 - p_2 \geq 0, \qquad 1 \geq x \geq 0, \tag{85}$$

is a monotonically decreasing function of temperature, as follows from Eq. (22). As seen from (83, 84), both

$\delta w^2_{\max}$ and $\delta w^2_{\rm av}$ are monotonically increasing functions of temperature T. It is obvious that $\delta w^2_{\max} > \delta w^2_{\rm av}$, except for the zero temperature situation $x = 1$, where they are both equal to zero. For very high temperatures, that is, for $x \to 0$, $\delta w^2_{\rm av} = 1/3$ in agreement with (81). Note that off-diagonal elements of $\hat{\Omega}$ increase $\delta w^2_{\max}$, while $\delta w^2_{\rm av}$ does not depend on them at all.

V. THERE IS NO DIRECT ANALOG OF THE CLASSICAL BK EQUALITY IN THE QUANTUM SITUATION.

The discussion in section IV B provides a definite evidence to think that in contrast to the classical case, the fluctuations of work in the quantum situation are not controlled by any *direct* analog of the classic BK equality (11). In the present section we give another illustration of this fact.

Assume for concreteness that the Gibbsian density matrix $\hat{\rho}$ in (21) was separated into pure subensembles by means of the measurement of $\hat{H}$, that is, the subensembles are described by pure density matrices $\{|\varepsilon_l\rangle\langle\varepsilon_l|\}_{l=1}^n$, where $\{|\varepsilon_l\rangle\}_{l=1}^n$ are eigenvectors of $\hat{\rho}$.

According to (39) one has for realizations of the random quantity work

$$w_l = \langle\varepsilon_l|\hat{U}_\tau^\dagger \hat{H}\hat{U}_\tau|\varepsilon_l\rangle - \varepsilon_l, \quad (86)$$

$$= \sum_{k=1}^n C_{kl}\varepsilon_k - \varepsilon_l, \quad l = 1, ..., n, \quad (87)$$

where

$$C_{kl} = |\langle\varepsilon_k|\hat{U}_\tau|\varepsilon_l\rangle|^2, \quad (88)$$

is a double-stochastic matrix:

$$\sum_{k=1}^n C_{kl} = \sum_{l=1}^n C_{kl} = 1. \quad (89)$$

Each of realizations w_l has probability p_l, as given by (22).

One now constructs

$$\langle e^{-\beta w}\rangle \equiv \sum_{l=1}^n p_l e^{-\beta w_l} = \frac{1}{Z}\sum_{l=1}^n e^{-\beta\sum_{k=1}^n C_{kl}\varepsilon_k}, \quad (90)$$

that is, averages $e^{-\beta w}$ directly as was done in the classical situation. As shown in Appendix B,

$$1 - \frac{\beta^2\Delta}{2Z}\, e^{-\beta\varepsilon_{\min}} \le \langle e^{-\beta w}\rangle \le 1 - \frac{\beta^2\Delta}{2Z}\, e^{-\beta\varepsilon_{\max}}, \quad (91)$$

$$\Delta \equiv \varepsilon^{\rm T}(1 - CC^{\rm T})\varepsilon$$

$$= \sum_{k=1}^n \left[\langle\varepsilon_k|\hat{H}|\varepsilon_k\rangle^2 - \langle\varepsilon_k|\hat{U}_\tau^\dagger\, \hat{H}\, \hat{U}_\tau|\varepsilon_k\rangle^2\right], \quad (92)$$

where $\varepsilon^{\rm T} = (\varepsilon_1, ..., \varepsilon_n)$ is the vector of eigenvalues of $\hat{H}$, Z is the partition sum defined in (21), and where $\varepsilon_{\min}$ and $\varepsilon_{\max}$ are the minimal and maximal ones among $(\varepsilon_1, ..., \varepsilon_n)$.

Since all the eigenvalues ν of the product of a double-stochastic matrix to its transpose satisfy $0 \le \nu \le 1$ [16], one has

$$\varepsilon^{\rm T}(1 - CC^{\rm T})\varepsilon \ge 0. \quad (93)$$

Thus $\langle e^{-\beta w}\rangle$ is strictly smaller than unity. As compared to our discussion of the classical situation in section II E, the result $\langle e^{-\beta w}\rangle < 1$ does not in general permit to draw quantum analogs of the classical features b (active realizations) and c (dispersion at high T) in section II E.

VI. COMPARISON WITH OTHER APPROACHES.

In the present section we study two approaches known in literature. The purpose is to understand whether they have the proper physical meaning for describing fluctuations of work. Since they both allow to generalize the classical BK equality (though in different ways), the adoption of either of them will mean —as we discuss in detail below— that there is no major qualitative difference in behavior of quantum and classical fluctuations of work. It should perhaps be stressed that our concern is the applicability of these approaches for describing fluctuations of work under conditions formulated in the Introduction; their usefulness for other purposes is neither discussed, nor criticized.

A. Observable of work.

Recall from definitions (29, 30) that for any initial ensemble described by $\hat{\rho}$, the average of $\hat{\Omega}$ is equal to the work done on the corresponding ensemble.

The approach goes on by stating [6, 13, 27, 32] that the operator $\hat{\Omega}$ is the "observable of work" in the standard sense of quantum observables [17], e.g., the quantity $\mathrm{tr}[\hat{\rho}\,\hat{\Omega}^2] - [\mathrm{tr}(\hat{\rho}\,\hat{\Omega})]^2$ is to be interpreted as the dispersion of work for any $\hat{\rho}$. However, while $\mathrm{tr}\,[\hat{\Omega}\,\hat{\rho}]$ happens to be

[16] For any double-stochastic matrix C_{ik}, consider the matrix $CC^{\rm T}$, where $C^{\rm T}$ is the transpose of C, and let a_i be an eigenvector of $CC^{\rm T}$ corresponding to a (necessarily non-negative) eigenvalue ν: $\sum_{k,l=1}^n C_{ik}C_{lk}a_l = \nu a_i$. One has $|\sum_{k=1}^n C_{ik}a_k| = |\nu a_i| = \nu\,|a_i| \le \sum_{k,l=1}^n C_{ik}C_{lk}|a_l|$, and then $\nu\sum_{i=1}^n |a_k| \le \sum_{k=1}^n |a_k|$, that is, $\nu \le 1$.

[17] Once $\hat{\Omega}$ is given an independent meaning as a quantum observable, there arises a question on its measurability, since the standard theories of quantum measurements, see e.g. [46, 47], operate in Schrödinger representation. We shall not pursue this problem here, but rather take as working hypothesis that this measurement can be carried out.

equal to the average energy lost by the work source $\mathcal{W}$, simply due to conservation of the average energy during the system-work-source interaction, this alone is, of course, not sufficient to regard $\hat{\Omega}$ as an operator of work. In fact, such an interpretation relies on the analogy between the definition (30) of $\hat{\Omega}$ and the classical expression (8) for energy difference. Such analogies are very widespread in general, and once it is accepted that $\hat{\Omega}$ represents the proper energy difference operator, the extension of its interpretation toward operator of work seems rather natural.

Let us however recall from our discussion in the Introduction that we expect for a proper approach to fluctuations of work to apply in arbitrary non-equilibrium situation. It is now possible to argue that in general $\hat{\Omega}$ does not have the proper meaning of energy difference operator, let alone its meaning as the operator of work.

Let the ensemble $\mathcal{E}(\hat{\rho})$ have a density matrix $\hat{\rho}(0) = |0\rangle\langle 0|$, such that $|0\rangle$ is an eigenstate of $\hat{\Omega} \equiv \hat{U}_\tau^\dagger \, \hat{H} \, \hat{U}_\tau - \hat{H}$ with eigenvalue zero:

$$\hat{\Omega}|0\rangle = 0. \tag{94}$$

Recall that $\hat{U}_\tau^\dagger \, \hat{H}(\tau) \, \hat{U}_\tau$ is the Hamiltonian in the Heisenberg representation a time τ, while the Schrödinger picture relation $\hat{H}(\tau) = \hat{H}$ is due to the assumed cyclic feature of the process.

In general,

$$[\hat{U}_\tau^\dagger \, \hat{H} \, \hat{U}_\tau, \hat{H}] \neq 0, \tag{95}$$

so that $|0\rangle$ is neither an eigenstate of $\hat{U}_\tau^\dagger \, \hat{H} \, \hat{U}_\tau$, nor an eigenstate of $\hat{H}$.

According to quantum mechanics, Esq. (94) should be interpreted as follows: the operator $\hat{\Omega}$ has on the ensemble $\mathcal{E}(|0\rangle\langle 0|)$ a definite value equal to zero, that is, if it is interpreted as the operator of energy change, then for *all single systems* from $\mathcal{E}(|0\rangle\langle 0|)$ the energy does not change during this thermally isolated process.

There are however concrete examples —see Appendix I— showing that (94) can be consistent with

$$\left\langle 0 \,\middle|\, \left[\hat{U}_\tau^\dagger \, \hat{H} \, \hat{U}_\tau\right]^m \,\middle|\, 0 \right\rangle \neq \langle 0|\hat{H}^m|0\rangle, \quad \text{for} \quad m > 2. \tag{96}$$

This shows that the energy does change, since some of its moments do. In other words, the interpretation of $\hat{\Omega}$ as the energy difference operator is in general unsupportable. Note that the non-commutativity feature as expressed by (95) is essential for this conclusion.

1. *Restricted interpretation of* $\hat{\Omega}$.

A more restricted interpretation of $\hat{\Omega}$ can be given in the light of the definition of fluctuations of work discussed in section III. This will also show that if ρ commutes with $\hat{\Omega}$ (a semiclassical assumption), our approach is consistent with that of the observable of work.

Let the eigenresolution of $\hat{\Omega}$ be

$$\hat{\Omega} = \sum_{k=1}^{n} \omega_k |\omega_k\rangle\langle\omega_k|. \tag{97}$$

Note that for $\hat{\Omega}$ to have the meaning of the operator of work it is necessary that *i)* its eigenvalues $\{\omega_k\}_{i=1}^n$ have the meaning of work by themselves, i.e., ω_k should have both the meaning of average energy lost by the work source $\mathcal{W}$ and the average energy gained by a quantum ensemble, as we discussed in section III D; *ii)* probabilities of these realizations of work done on the initial ensemble $\mathcal{E}(\hat{\rho})$ should be given as $\{\, \langle\omega_k|\hat{\rho}|\omega_k\rangle \,\}_{k=1}^n$.

Now, if $\hat{\rho}$ and $\hat{\Omega}$ commute,

$$[\hat{\rho}, \hat{\Omega}] = 0, \tag{98}$$

then their eigenvectors can be chosen the same, and, by measuring $\hat{\Omega}$, $\hat{\rho} = \sum_{k=1}^n p_k |\omega_k\rangle\langle\omega_k|$ can be separated into subensembles $\{\mathcal{E}(|\omega_k\rangle\langle\omega_k|)\}_{k=1}^n$ with probabilities $p_k = \langle\omega_k|\hat{\rho}|\omega_k\rangle$. The average work done on each subensemble $\mathcal{E}(|\omega_k\rangle\langle\omega_k|)$ is then equal to $\omega_k = \langle\omega_k|\hat{\Omega}|\omega_k\rangle$, and one can admit the restricted interpretation of $\hat{\Omega}$ as an operator of work.

Conversely, if $\hat{\rho}$ can be separated into subensembles,

$$\hat{\rho} = \sum_{k=1}^{n} \lambda_k |\psi_k\rangle\langle\psi_k|, \tag{99}$$

and if each of them is let to interact with the work source $\mathcal{W}$ such that

$$\omega_k = \langle\psi_k|\hat{\Omega}|\psi_k\rangle, \quad \lambda_k = \langle\omega_k|\hat{\rho}|\omega_k\rangle, \tag{100}$$

then three conditions (99, 100, 97) imply commutation (98).

To show this we proceed in a slightly indirect way, which is useful by itself. It can be noted that the dispersion

$$\mathrm{tr}[\hat{\rho}\,\hat{\Omega}^2] - [\mathrm{tr}(\hat{\rho}\,\hat{\Omega})]^2 = \sum_{k=1}^{n} \langle\omega_k|\hat{\rho}|\omega_k\rangle \, (\omega_k - W)^2 \tag{101}$$

of the operator $\hat{\Omega}$ provides an upper bound for the maximal dispersion $\delta w_{\max}^2$ of work given by Eq. (60):

$$\begin{aligned} &\mathrm{tr}(\hat{\rho}\hat{\Omega}^2) - W^2 - \delta w_{\max}^2 \\ &= \frac{1}{2} \sum_{i,k=1}^{n} \frac{(p_i - p_k)^2}{p_i + p_k} \, |\langle p_k|\hat{\Omega}|p_i\rangle|^2 \geq 0. \end{aligned} \tag{102}$$

The equality in the RHS of (102) is realized only if $\hat{\rho}$ and $\hat{\Omega}$ commute, that is, either $\langle p_k|\hat{\Omega}|p_i\rangle$ is zero for $i \neq k$, or for some pair $i \neq k$ one has $\langle p_k|\hat{\Omega}|p_i\rangle \neq 0$, but the corresponding eigenvalues of $\hat{\rho}$ are degenerate: $p_i = p_k$. Thus $\delta w_{\max}^2$ can be equal to $\mathrm{tr}(\hat{\rho}\hat{\Omega}^2) - W^2$ only if $[\hat{\rho}, \hat{\Omega}] = 0$.

Now note that if Eq. (99, 100, 97) are assumed to be valid, they imply $\mathrm{tr}(\hat{\rho}\hat{\Omega}^2) - W^2 - \delta w^2_{\max} \leq 0$ simply due to the definition of the maximal dispersion. This is consistent with Eq. (102) only for $\mathrm{tr}(\hat{\rho}\hat{\Omega}^2) - W^2 - \delta w^2_{\max} = 0$, which implies $[\hat{\rho}, \hat{\Omega}] = 0$ as we saw above. We conclude that (99, 100, 97) imply (98), as was promised.

Thus, when $[\hat{\rho}, \hat{\Omega}] \neq 0$, $\hat{\Omega}$ does not qualify as the operator of work even in the restricted sense. We also conclude that though the approach does predict an upper bound for δw^2, this bound is not reachable [18].

2. On a generalization of the classical BK equality.

Though $\hat{\Omega}$ does not have the meaning of the operator of work — except in the restricted sense and under condition (98) — there is an operator generalization of Eqs. (11, 12) which was proposed by Bochkov and Kuzovlev in [13, 15]:

$$\frac{\mathrm{tr}\, e^{-\beta\hat{\Omega}-\beta\hat{H}}}{Z} = \left\langle \overrightarrow{\exp}\left[-\int_0^\beta \mathrm{d}s\, e^{-s\hat{H}}\hat{\Omega}\, e^{s\hat{H}} \right] \right\rangle \quad (103)$$

$$\equiv \mathrm{tr}\left(\overrightarrow{\exp}\left[-\int_0^\beta \mathrm{d}s\, e^{-s\hat{H}}\hat{\Omega}\, e^{s\hat{H}} \right] \hat{\rho} \right) = 1. \quad (104)$$

We recall its derivation in Appendix E [19]. A similar relation was derived in [27].

Let us work out some consequences of (104). As compared to the classic case, the matters are complicated by the presence of anti-time-ordering and the integral $\int_0^\beta$ in (103, 104). If one would insist on not having them, then the equality (103, 104) can still be converted into an inequality. By applying Thompson-Golden inequality [55] [20], $\mathrm{tr}\,[\, e^{\hat{A}}\, e^{\hat{B}}\,] \geq \mathrm{tr}\, e^{\hat{A}+\hat{B}}$, valid for any hermitian operators $\hat{A}$ and $\hat{B}$ (the equality sign is realized here if and only if $[\hat{A}, \hat{B}] = 0$), one gets

$$\langle e^{-\beta\hat{\Omega}} \rangle \equiv \mathrm{tr}\,[\, \hat{\rho}\, e^{-\beta\hat{\Omega}}\,]$$

$$= \sum_{k=1}^{n} \langle \omega_k|\hat{\rho}|\omega_k\rangle\, e^{-\beta\omega_k} \geq \frac{1}{Z}\,\mathrm{tr}\, e^{-\beta\hat{\Omega}-\beta\hat{H}} = 1, \quad (105)$$

where $|\omega_k\rangle$ and ω_k are eigenvectors and eigenvalues of $\hat{\Omega}$ as defined by (97).

If now we could interpret $\hat{\Omega}$ as the operator of work, that is, if the eigenvalues ω_k of $\hat{\Omega}$ would have the meaning of work by themselves, we would note that $\langle \omega_k|\hat{\rho}|\omega_k\rangle$ is the probability of observing the eigenvalue ω_k upon the measurement of $\hat{\Omega}$ on the state $\hat{\rho}$, and then Eq. (105) would allow us to study fluctuations of work exactly in the way we did in section II E for the classical situation. We would then draw the same general conclusions, and the fact that (105) is an inequality will only *strengthen* these conclusions as compared to the classical situation. However, as we saw above, it is impossible to identify $\hat{\Omega}$ with the operator work, and thus fluctuations of work cannot be studied on the base of (105), except for $[\hat{\rho}, \hat{\Omega}] = 0$, where Eqs. (103, 104, 105) reduce to the usual (essentially classical) BK equality.

B. On the approach based on two-time measurements of energy.

Yet another, different approach to fluctuations of work and extension of the classical BK equality was proposed in Refs. [28, 29, 30]. We shall present it in a more extended form, since it is necessary for the understanding of its proper physical meaning. On the other hand, in order do not dwell into unnecessary technical details, we shall assume that the spectrum of the Hamiltonian $\hat{H}$ is non-degenerate (compare with (23))

$$\varepsilon_1 < \varepsilon_2 < ... < \varepsilon_n. \quad (106)$$

At the time $t = 0$ one measures energy (corresponding to the operator $\hat{H}$) for the ensemble described by the gibbsian density matrix (21). The probability to get an eigenvalue ε_l of $\hat{H}$ is seen from (20) to be

$$p(l|\mathcal{M}_0) = \langle \varepsilon_l|\hat{\rho}|\varepsilon_l\rangle, \quad (107)$$

$$= p_l. \quad (108)$$

Eq. (107) is the general quantum formula (Born's rule), while Eq. (108) follows from the Gibbsian form (21, 19, 22) of $\hat{\rho}$. The symbol $\mathcal{M}_0$ in (107) reminds that the probability is conditional and refers to the measurement of $\hat{H}$ done at $t = 0$. The necessity of such explicit notations will be seen below. Formally it is always allowed, since *any* probability is conditional.

According to Wigner's formula for multi-time probabilities in quantum mechanics [49], the subsequent measurement of energy at the time τ — represented by the same Hamiltonian $\hat{H}$ due to the cyclic feature of the considered process — will then produce a result ε_k with the conditional probability

$$p(k|l, \mathcal{M}_\tau, \mathcal{M}_0) = |\langle \varepsilon_k|\hat{U}_\tau|\varepsilon_l\rangle|^2. \quad (109)$$

There three conditionals for the probability in the LHS of (109): $\mathcal{M}_0$ and $\mathcal{M}_\tau$ stand for the measurements done

---

[18] Note that the difference $\mathrm{tr}(\hat{\rho}\hat{\Omega}^2) - W^2 - \delta W^2 = \sum_{\alpha=1}^N \lambda_\alpha \left(\langle\psi_\alpha|\hat{\Omega}^2|\psi_\alpha\rangle - \langle\psi_\alpha|\hat{\Omega}|\psi_\alpha\rangle^2 \right) \geq 0$ is by itself always non-negative for any separation of $\hat{\rho}$ into subensembles.

[19] For the equilibrium ensemble (21), the Thomson formulation of the second law can be derived from (103, 104) upon the application of the Peierls-Bogoliubov inequality (recalled in Appendix E): $e^{-\beta\,\mathrm{tr}[\hat{\rho}\,\hat{\Omega}]} \leq \frac{1}{Z}\,\mathrm{tr}\, e^{-\beta\hat{\Omega}-\beta\hat{H}} = 1$. From this it follows once again that $W = \mathrm{tr}\,[\hat{\Omega}\,\hat{\rho}] \geq 0$.

[20] Thompson-Golden inequality is a particular consequence of the following submajorization relation $\lambda\left(e^{\hat{A}+\hat{B}}\right) \prec_w \lambda\left(e^{\hat{A}/2}e^{\hat{B}}e^{\hat{A}/2}\right)$, where $\lambda\left(\hat{A}\right)$ is the eigenvalue vector of a hermitian operator $\hat{A}$; see [58] for more details.

at $t = 0$ and $t = \tau > 0$, while the index l indicate on the result ε_l got during the first measurement. The meaning of (109) is that the ensemble of systems which during the first measurement at $t = 0$ produced the result ε_l, is described for $t > 0$ by $|\varepsilon_l\rangle\langle\varepsilon_l|$. The members of this ensemble couple to the work source $\mathcal{W}$, the state evolves to $\hat{U}_\tau\,|\varepsilon_l\rangle\langle\varepsilon_l|\,\hat{U}_\tau^\dagger$ at the time $t = \tau$, and then is subjected to the second measurement.

Thus the total probability for having the result ε_l at the moment $t = 0$ *and* the result ε_k at $t = \tau$ is given by

$$p(k,l|\mathcal{M}_0,\mathcal{M}_\tau) = p(l/\mathcal{M}_0,\mathcal{M}_\tau)\,p(k|l,\mathcal{M}_\tau,\mathcal{M}_0) \quad (110)$$
$$= p(l|\mathcal{M}_0)\,p(k|l,\mathcal{M}_\tau,\mathcal{M}_0). \quad (111)$$

When passing from (110) to (111), we used the obvious relation $p(l|\mathcal{M}_0,\mathcal{M}_\tau) = p(l|\mathcal{M}_0)$ (no dependence on the future).

It is to be noted that

$$p(k|\mathcal{M}_0,\mathcal{M}_\tau) = \sum_{l=1}^{n} p(k,l|\mathcal{M}_0,\mathcal{M}_\tau)$$
$$= \sum_l p_l \langle\varepsilon_k|\hat{U}_\tau\,|\varepsilon_l\rangle\langle\varepsilon_l|\,U_\tau^\dagger|\varepsilon_k\rangle, \quad (112)$$

that is, the probability to have the result ε_k at the second measurement is for a general initial density matrix $\hat{\rho}$ not equal to

$$p(k|\mathcal{M}_\tau) = \langle\varepsilon_k|\hat{U}_\tau\,\hat{\rho}\,U_\tau^\dagger|\varepsilon_k\rangle, \quad (113)$$

which is the probability to get the result k in a different context, where no first measurement was done. Such an equality is valid, though, if $\hat{\rho}$ commutes with $\hat{H}$, which is the case with the Gibbsian density matrix (21). Let us first restrict our attention to this case. One notes from (109) the double-stochastic feature of $p(k|l,\mathcal{M}_\tau,\mathcal{M}_0)$:

$$\sum_{k=1}^{n} p(k|l,\mathcal{M}_\tau,\mathcal{M}_0) = \sum_{l=1}^{n} p(k|l,\mathcal{M}_\tau,\mathcal{M}_0) = 1, \quad (114)$$

and calculates using (19, 22, 108, 114):

$$\langle e^{-\beta(\varepsilon_k-\varepsilon_l)}\rangle_{0,\tau}$$
$$\equiv \sum_{k,l=1}^{n} p(l|\mathcal{M}_0)\,p(k|l,\mathcal{M}_\tau,\mathcal{M}_0)\,e^{-\beta(\varepsilon_k-\varepsilon_l)}$$
$$= \frac{1}{Z}\sum_{k,l=1}^{n} p(k|l,\mathcal{M}_\tau,\mathcal{M}_0)\,e^{-\beta\varepsilon_k} = 1. \quad (115)$$

This is the equality got in Refs. [28, 29, 30] as a generalization of the classic BK equality.

Note that for the density matrix (21) the average

$$\sum_{k,l=1}^{n} p(l|\mathcal{M}_0)\,p(k|l,\mathcal{M}_\tau,\mathcal{M}_0)\,(\varepsilon_k-\varepsilon_l) = W, \quad (116)$$

is equal to the work as defined by (29). The statement of the second law, $W \geq 0$, can once again be deduced from (115) by employing convexity of the exponent.

1. *Critique of the approach.*

Would now we be able to associate the work with a random variable having realizations $\{\varepsilon_k - \varepsilon_l\}_{k,l=1}^n$ and the corresponding probabilities $\{p(k|l,\mathcal{M}_\tau,\mathcal{M}_0)\}_{k,l=1}^n$, it would be possible to study fluctuations of work on the base of Eq. (115), and to draw essentially the same conclusions as we did in section II E for the classical case. It is, however, not difficult to see that the same criticisms we brought in section VI A with respect to the "observable of work" applies here too.

Keeping in mind our discussion after Eq. (109), note that if the ensemble initially described by $|\varepsilon_l\rangle\langle\varepsilon_l|$ couples to the work source $\mathcal{W}$, its mechanical degree of freedom looses at the time $t = \tau$ the energy

$$\mathrm{tr}\left(\hat{\Omega}\,|\varepsilon_l\rangle\langle\varepsilon_l|\right) = \mathrm{tr}\left(\hat{H}\,\hat{U}_\tau\,|\varepsilon_l\rangle\langle\varepsilon_l|\,\hat{U}_\tau^\dagger\right) - \varepsilon_l. \quad (117)$$

Since the final density matrix $\hat{U}_\tau\,|\varepsilon_l\rangle\langle\varepsilon_l|\,\hat{U}_\tau^\dagger$ need not commute with $\hat{H}$, the energy need not any definite value at that time, and Eq. (117) does in general not reduce to $\varepsilon_k - \varepsilon_l$ with any k. Such a reduction takes place when

$$\mathrm{tr}\left(\hat{H}\,\hat{U}_\tau\,|\varepsilon_l\rangle\langle\varepsilon_l|\,\hat{U}_\tau^\dagger\right) = \sum_{k=1}^{n} C_{kl}\varepsilon_k = \varepsilon_{\pi(l)}, \quad (118)$$
$$\text{for} \quad l = 1,...,n, \quad (119)$$

where C_{kl} is defined via (88), and where $(\pi(1),...,\pi(n))$ is some permutation of the sequence $(1,...,n)$. Eq. (118) can now be re-written as

$$\sum_{k=1}^{n} \widetilde{C}_{kl}(\varepsilon_k - \varepsilon_l) = 0, \quad (120)$$

where the matrix $\widetilde{C} = C\,\Pi$ the product of C and the corresponding permutation matrix Π, and where we noted that once the matrices C and Π are double-stochastic (see (89) for definition), so is $\widetilde{C}$. Note with help of Eq. (106) that for $l = n$ all terms with $k \neq n$ in (120) are negative unless $\widetilde{C}_{k\neq n} = 0$, which via the double-stochastic feature of $\widetilde{C}$ implies: $\widetilde{C}_{n\neq k} = 0$ and $\widetilde{C}_{nn} = 1$. Continuing along the same lines for $l < n$, one gets that (118) can take place only when $\widetilde{C}$ reduces to unity matrix, or, equivalently, C reduces to a permutation matrix:

$$\hat{U}_\tau\,|\varepsilon_l\rangle\langle\varepsilon_l|\,\hat{U}_\tau^\dagger = |\varepsilon_{\pi(l)}\rangle\langle\varepsilon_{\pi(l)}|. \quad (121)$$

Thus, in general it is (117) and not $\varepsilon_k - \varepsilon_l$ that can be interpreted as the work for this single realization occuring with probability p_l, and this is precisely the point from which we departed in section III.

It is also straightforward to see that the approach does not apply out of equilibrium. The reasons for this are more straightforward than for the previous approach.

Recall from the Introduction that the proper definition of fluctuations of work is expected to apply to any non-equilibrium initial ensemble $\mathcal{E}(\hat{\sigma})$ with density matrix $\hat{\sigma}$

not commuting with $\hat{H}$:

$$[\hat{\sigma}, \hat{H}] \neq 0. \tag{122}$$

In particular, the work averaged over those fluctuations should be equal to the one done on the ensemble.

The present approach is generalized uniquely for arbitrary initial state: the definitions of $p(l|\mathcal{M}_0)$ and $p(k|l, \mathcal{M}_\tau, \mathcal{M}_0)$ in Eqs. (109, 107) remains unaltered: one substitutes there $\hat{\sigma}$ instead of $\hat{\rho}$.

It is now straightforward to see from (122) that due to non-diagonal terms in $\hat{\sigma}$, the average $\sum_{k.l=1}^{n} p(l|\mathcal{M}_0)\, p(k|l, \mathcal{M}_\tau, \mathcal{M}_0)\, (\varepsilon_k - \varepsilon_l)$ is not equal to the work $\mathrm{tr}(\hat{\sigma}\hat{\Omega})$ done on the overall ensemble:

$$\mathrm{tr}(\hat{\sigma}\hat{\Omega}) - \sum_{k.l=1}^{n} p(l|\mathcal{M}_0)\, p(k|l, \mathcal{M}_\tau, \mathcal{M}_0)\, (\varepsilon_k - \varepsilon_l)$$
$$= \mathrm{tr}\left[\hat{U}_\tau^\dagger \hat{H} \hat{U}\, (\, \hat{\sigma} - |\varepsilon_l\rangle\langle\varepsilon_l|\hat{\sigma}|\varepsilon_l\rangle\langle\varepsilon_l|\,)\right] \neq 0. \tag{123}$$

2. The approaches based on the "observable of work" and on two-time measurements of energy are different.

This is seen already by comparing Eq. (105) with (115). Still we want to understand this difference in more detail. More precisely, though for the initial density matrix commuting with $\hat{H}$, the first and the second moments generated by the two approaches are equal:

$$\mathrm{tr}[\hat{\rho}\hat{\Omega}^p] = \sum_{k,l=1}^{n} p(l|\mathcal{M}_0)\, p(k|l, \mathcal{M}_\tau, \mathcal{M}_0)\, (\varepsilon_k - \varepsilon_l)^p, \tag{124}$$

$$p = 0, 1, 2, \tag{125}$$

already the third moments are in general different, even for $[\hat{\rho}, \hat{H}] = 0$. Indeed, assuming validity of the latter condition, one gets

$$\mathrm{tr}[\hat{\rho}\hat{\Omega}^3] - \sum_{k,l=1}^{n} p(l|\mathcal{M}_0)\, p(k|l, \mathcal{M}_\tau, \mathcal{M}_0)\, (\varepsilon_k - \varepsilon_l)^3$$
$$= \mathrm{tr}\left(\hat{\Omega}\,[\hat{\Omega}, \hat{\rho}]\,\hat{H}\right) = \mathrm{tr}\left(\hat{\rho}\,[\hat{H}, \hat{\Omega}]\,\hat{\Omega}\right). \tag{126}$$

The RHS of (126) vanishes if $[\hat{\rho}, \hat{\Omega}] = 0$, or $[\hat{H}, \hat{\Omega}] = 0$, in addition to $[\hat{\rho}, \hat{H}] = 0$.

Note as an illustration that for the two-level example of section IV D the RHS of (126) reads:

$$\mathrm{tr}\left(\hat{\rho}\,[\hat{H}, \hat{\Omega}]\,\hat{\Omega}\right) = (p_1 - p_2)(\varepsilon_1 - \varepsilon_2)|\chi|^2, \tag{127}$$

where $\hat{\Omega}$ is given by (82), and where p_k and ε_k are eigenvalues of $\hat{\rho}$ and $\hat{H}$, respectively. For the Gibbsian density matrix $\hat{\rho}$, the RHS of (127) has negative sign.

Finally note that differences between the two approaches were recently studied in [32] in a different context.

C. Summary.

We have discussed two approaches known in literature, and argued that in the proper quantum domain, though being related to energy, they do not describe fluctuations of work. Work is a rather particular form of energy having several specific features we discussed in section II B. These two approaches allow different generalizations of the classical BK equality and this makes them operationally close to the classical situation. Still this resemblance is superficial, since, as we argued, it is not ensured by the two approaches that realizations of the claimed random quantity/operator of work have themselves the physical meaning of work.

VII. CONCLUSION.

The second law has a statistical character as it is both formulated and valid for ensembles of identically prepared systems. It is therefore of interest to investigate this statistical aspect in more detail. For the entropic formulation of the second law, this analysis is by now a standard chapter of statistical thermodynamics [1, 2, 3].

In the present paper we studied how Thomson's formulation of the second law: no work from an equilibrium ensemble by a cyclic process, emerges through averaging over fluctuations of work in the quantum situation. It will be useful at this moment to recall the special role of Thomson's formulation, and then to proceed with concluding remarks on our results.

1. The main features of Thomson's formulation of the second law.

a. The formulation uses the concept of work which is unambiguously defined both conceptually and operationally, both in and out of equilibrium. In this respect work is contrasting to entropy, which is well-defined only in (nearly) equilibrium states of macroscopic systems.

b. Thomson's formulation is valid for any *finite* [21] or *infinite* [17], quantum or classical system interacting with macroscopic sources of work. In this context one notes that not all formulations of the second law have such a universal regime of validity. While all formulations

[21] In this context one sometimes hears that the second law must refer to macroscopic systems, and there is no sense in applying it for finite systems. This opinion is not correct, as instanced by Thomson's formulation. Would it not be valid for a finite system coupled to work sources, the very its application to macroscopic systems will be endangered, because the initial Gibbsian ensemble (21) is prepared under weak interaction with an equilibrium thermal bath, see e.g. [4]: any cycle violating the formulation for a finite system can be repeated to achieve a violation for the bath.

are equivalent in the standard thermodynamical domain, that is, for (nearly) equilibrium states of macroscopic systems, some of them have definite limits when considered for finite systems [20] or for low temperatures (quantum domain) [11, 12].

c. In its literal form Thomson's formulation does not imply any irreversibility, since the dynamics of the system coupled to work source is unitary and thus formally reversible: if some work was put into the initially equilibrium system it can in principle be extracted back. The irreversibility with respect to work transfer comes into existence when one takes into account that in practice no work source can interact with all possible degrees of freedom. In particular, if the system was subjected to a thermal bath after it had interacted with the work source, it relaxes back to its Gibbsian state and the work which had been put into it cannot be recovered by *any* work-source acting on the system only (a similar argument is presented in [6], chapter 5). Thus the relation between Thomson's formulation and irreversibility is seen clearly at least in the simplest situation.

d. It should perhaps be stressed that Thomson's formulation does not refer to all aspects usually associated with the second law, e.g., by itself it does not explain how a subsystem of a proper macroscopic system (thermal bath) relaxes toward a Gibbsian equilibrium state (though on the basis of Thomson's formulation it is still possible to argue that – under certain assumptions – the Gibbsian state is the only one which forbids work extraction via *any* cyclic thermally isolated process [17, 18]). The property of relaxation toward a Gibbsian state is to be viewed as an independent item of statistical physics; its standard classical understanding was reshaped in literature various times; see, e.g., [11, 12, 36, 37].

2. *What appeared to be problematic in defining fluctuations of work in the quantum situation.*

As we saw in section VI, due to non-commutativity of various quantum observables, there are different quantum quantities which, in the classical limit, coincide with the random quantity work. Thus, as often, classical reasoning alone is of no help for defining fluctuations of work.

One therefore has first to state what basic features the fluctuations of work are going to have, as we did in section I, and then to recognize that the work is an essentially mechanical, classical quantity — in spite of the fact that it can be added or extracted from a quantum system — since it is an energy transferred to macroscopic degrees of freedom of the work-source that, at least in principle, should be retransferrable to other classical sources.

Once this feature was recognized, the definition of fluctuations of work we presented in section III is straightforward.

3. *What is similar and what is different in classical and quantum definitions of fluctuations of work?*

In both situations the definition of work as a random quantity employs the same idea: the initial homogeneous ensemble of identically prepared systems is separated into irreducible (homogeneous) subensembles. Both in quantum and classics these irreducible subensembles are described maximally completely. In classics they correspond to a trivial subensemble of identical copies of the same system (so that within a subensemble no fluctuations are present whatsoever), and they are described via phase-space points and trajectories. In quantum mechanics these subensembles, described by pure density matrices (wave functions), provide definite (non-fluctuating) values for the largest possible, *but non-exhaustive*, set of observables.

In classics the irreducible subensembles of the initial ensemble obviously exist *a priori*, that is, without need of any measurement. In the quantum situation the very structure of subensembles does depend on the measurement applied for the actual separation, or, in other words for the preparation of an ihomogeneous ensemble. As the main consequence, the separation of a mixed ensemble is not unique, and thus the random quantity work is *contextual* in the quantum situation. It is therefore to be recalled that in the quantum situation the definition of fluctuations of work always needs this initial *preparational* measurement, a step which is absent in classics.

In the second step, systems from each irreducible subensemble interact with the same macroscopic source which does on them the same thermally isolated process. Realizations of the random quantity work are then defined as the average energy increase of the work-source when interacting with each subensemble, while the probability of each realization is given by the weight of the corresponding subensemble in the initial mixed ensemble.

In this way the full physical meaning of work is kept, and the approach can be applied to any non-equilibrium initial state.

4. *Dispersion of work.*

The most direct quantity which characterizes fluctuations of work is the dispersion of work we studied in section IV. Although the work is a contextual random quantity and depends on the measurement that was done to separate the initial mixed ensemble into pure subensembles, one can define two reasonable quantities —maximal dispersion and the dispersion for a randomly chosen separation on the initial ensemble— that depend solely on the internal features of the system, that is, on its initial state and its time-dependent Hamiltonian. These quantities were calculated explicitly for any finite quantum system and studied in section IV.

5. Non-existence of the direct generalization of classic BK equality.

In the classical situation fluctuations of work in an initially equilibrium state are controlled by the BK equality [13, 21]. This equality allows to draw a number of model-independent statements on fluctuations of work which we summarized in section II E. In contrast, the *direct* generalization of BK equality to the quantum domain —the one which would allow to draw similar qualitative conclusion on fluctuations of work— does not exist; see section V. As we discussed in detail in section VI, there are quantum generalizations of BK equality, but they refer to quantities which describe fluctuations of work only if some classical features are present, e.g. those implied by Eq. (98). As the main consequence, fluctuations of work in the quantum situation can have features which are impossible in classics, e.g., (inter-subensemble) fluctuations can be absent completely.

Acknowledgments

It is a pleasure to thank Roger Balian for inspiring discussions. We thank Claudia Pombo for critical remarks and for stressing the importance of understanding of the concept of work.

The work of A.E. A is part of the research programme of the Stichting voor Fundamenteel Onderzoek der Materie (FOM, financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)).

[1] P.S. Epstein, *Textbook of Thermodynamics*, (Wiley & Sons, New York, 1937).
[2] R.C. Tolman, *The Principles of Statistical Mechanics*, (Oxford University Press, London, 1938).
[3] L.D. Landau and E.M. Lifshitz, *Statistical Physics, I*, (Pergamon Press Oxford, 1978).
[4] R. Balian, *From Microphysics to Macrophysics*, volume I, (Springer, 1992).
[5] F. Schlögl, *Probability and Heat*, (Vieweg, Braunschweig/Wiesbaden, 1989).
[6] G. Lindblad, *Non-Equilibrium Entropy and Irreversibility*, (D. Reidel, Dordrecht, 1983).
[7] J.C. Maxwell, Nature, **17**, 257 (1878).
[8] H.S. Leff and A.F. Rex, *Maxwell's Demon: Entropy, Information, Computing*, (Princeton: Princeton University Press, 1990).
[9] J.W. Gibbs, *Elementary Principles in Statistical Mechanics*, (New Haven, 1914).
[10] K. Sekimoto, F. Takagi, and T. Hondou Phys. Rev. E **62**, 7759 (2000).
[11] A.E. Allahverdyan and Th.M. Nieuwenhuizen, Phys. Rev. Lett **85**, 1799 (2000); J. Phys. A, **36**, 875 (2003); Phys. Rev. E **64** 056117 (2001).
[12] Th.M. Nieuwenhuizen and A.E. Allahverdyan, Phys. Rev. E **66**, 036102 (2002)
[13] G.N. Bochkov and Yu.E. Kuzovlev, Sov. Phys. JETP, **45**, 125 (1977).
[14] G.N. Bochkov and Yu.E. Kuzovlev, Sov. Phys. JETP, **49**, 543 (1979).
[15] G.N. Bochkov and Yu.E. Kuzovlev, Physica A **106**, 443 (1981); *ibid*, 480 (1981).
[16] I.M. Bassett, Phys. Rev. A **18**, 2356 (1978).
[17] W. Pusz and S.L. Woronowicz, Comm. Math. Phys., **58**, 273 (1978).
[18] A. Lenard, J. Stat. Phys., **19**, 575 (1978).
[19] W. Thirring, *A Course in Mathematical Physics 4: Quantum mechanics of large systems,* (Springer, Vienna, 1983).
[20] A.E. Allahverdyan and Th.M. Nieuwenhuizen, *Adiabatic processes need not correspond to optimal work*, cond-mat/0401548.
[21] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[22] G.E. Crooks, Phys. Rev. E **61**, 2361-2366 (2000); *ibid*, **60**, 2721 (1999).
[23] G.M. Wang, *et.al.*,, Phys. Rev. Lett., **89**, 050601 (2002).
[24] F. Ritort, Séminaire Poincaré, **2**, 63 (2003).
[25] A.E. Allahverdyan and Th.M. Nieuwenhuizen, cond-mat/0207587.
[26] C. Maes, Séminaire Poincaré, **2**, 29 (2003); also available at http://itf.fys.kuleuven.ac.be/christ
[27] S. Yukawa, J. Phys. Soc. Jpn. **69**, 2370 (2000); cond-mat/0007456.
[28] J. Kurchan, cond-mat/0007360.
[29] H. Tasaki, cond-mat/0009244.
[30] S. Mukamel, Phys. Rev. Lett. **90**, 170604 (2003).
[31] W. De Roeck and C. Maes, cond-mat/0309498.
[32] T. Monnai and S. Tasaki, cond-mat/0308337.
[33] P. Perrot, *A to Z of Thermodynamics*, (Oxford University Press, 1998).
[34] *Quantum Limits to the Second Law*, ed. D.P. Sheehan, (AIP Conf. Proc. 643, 2002)
[35] A.E. Allahverdyan and Th.M. Nieuwenhuizen, Physica A **305**, 542 (2002).
[36] P. Bocchieri and A. Loinger, Phys. Rev. **114**, 948 (1959)
[37] H. Tasaki, Phys. Rev. Lett. **80**, 1376 (1998).
[38] G. Laufer, Am. J. Phys. **51**, 42 (1983).
[39] A.J. Mallinckrodt and H.S. Leff, Am. J. Phys. **60**, 356 (1992).
[40] A.E. Allahverdyan, R. Balian and Th.M. Nieuwenhuizen, *Maximal work extraction from quantum systems*, cond-mat/0401574.
[41] A.E. Allahverdyan, R. Balian and Th.M. Nieuwenhuizen, cond-mat/0408316.
[42] A.E. Allahverdyan, R. Balian and Th.M. Nieuwenhuizen, *Defining work in quantum statistical mechanics*, in preparation.
[43] W.M. Elsasser, Phys. Rev. **52**, 987 (1937).
[44] J.L. Park, Am. J. Phys., **36**, 211 (1968). Philosophy of Science, **35**, 205 (1968).
[45] B. D'Espagnat, *Veiled Reality*, (Addison-Wesley, MA, 1995).
[46] A. Peres, *Quantum Theory: Concepts and Methods*, (Kluwer Academic Publishers, 1995).
[47] W.M. de Muynck, *Foundations of Quantum Mechanics,*

an Empirist Approach, (Kluwer Academic Publishers, 2002).
[48] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*, (Oxford University Press, Oxford, 2002).
[49] E.P. Wigner, Am. J. Phys., **31**, 6 (1963).
C.W. Gardiner, *Quantum Noise*, (Springer, 1996).
[50] A. Abragam and M. Goldman, *Nuclear Magnetism*, Clarendon Press, Oxford, 1982.
A. Abragam, *Principles of Nuclear Magnetism*, Clarendon Press, Oxford, 1961.
[51] E. Schrödinger, Proc. Cambridge Philos. Soc. **32**, 446 (1936).
[52] E. T. Jaynes, Phys. Rev. **108**, 171 (1957).
[53] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).
[54] A. Bassi and G. Ghirardi, Phys. Lett. A, **309**, 24 (2003).
[55] S. Golden, Phys. Rev. **137**, B1127 (1965).
C.J. Thompson, J. Math. Phys. **6**, 1812 (1965).
[56] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, (Academic Press, New York, 1979);
[57] P.M. Alberti and A. Uhlmann, *Stochasticity and Partial Order* (D. Reidel Publishing Co., 1982).
[58] R. Bhatia, *Matrix Analysis*, (Springer-Verlag, Berlin, 1997).
[59] T. Ando, Linear Algebra and its Applications, **118**, 163 (1989).
[60] D.S. Mitrinović, *Analytical Inequalities*, (Springer-Verlag, 1970).
[61] P. Kok and S. L. Braunstein, Phys. Rev. A **61**, 042304 (2000)
[62] C.A. Fuchs and C.M. Caves, Open Systems & Information Dynamics **3**, 1 (1995); quant-ph/9604001.
[63] M. Ozawa, quant-ph/0310072.
[64] O. Cohen, Phys. Rev. A **63**, 016102 (2001).
[65] D. R. Terno, Phys. Rev. A **63**, 016101 (2001).
[66] K.A. Kirkpatrick, quant-ph/0405058.
B. d'Espagnat, quant-ph/0111081.

APPENDIX A: DERIVATION OF EQ. (15).

Here we recall from [60] a generalization of Cauchy inequality used in Eq. (15).

Denote by $\Gamma = (x, p)$ the phase space point. Assume that al the integrals over the phase-space used below are finite. The desired inequality reads: if $a(\Gamma)$, $b(\Gamma)$, $x(\Gamma)$ are some functions satisfying

$$\int \mathrm{d}\Gamma\, a(\Gamma)\, x(\Gamma) = 0, \quad \int \mathrm{d}\Gamma\, b(\Gamma)\, x(\Gamma) = 1, \qquad \text{(A1)}$$

then

$$\int \mathrm{d}\Gamma\, x^2(\Gamma) \geq \frac{\int \mathrm{d}\Gamma\, a^2(\Gamma)}{\int \mathrm{d}\Gamma\, a^2(\Gamma) \int \mathrm{d}\Gamma\, b^2(\Gamma) - \left[\int \mathrm{d}\Gamma\, a(\Gamma)\, b(\Gamma)\right]^2}. \qquad \text{(A2)}$$

To prove this, define

$$A = \int \mathrm{d}\Gamma\, a^2(\Gamma), \qquad \text{(A3)}$$

$$B = \int \mathrm{d}\Gamma\, b^2(\Gamma), \qquad \text{(A4)}$$

$$C = \int \mathrm{d}\Gamma\, a(\Gamma)\, b(\Gamma), \qquad \text{(A5)}$$

$$y(\Gamma) = \frac{A\, b(\Gamma) - C\, a(\Gamma)}{AB - C^2}, \qquad \text{(A6)}$$

and note that

$$\int \mathrm{d}\Gamma\, x^2(\Gamma) \geq \int \mathrm{d}\Gamma\, y^2(\Gamma), \qquad \text{(A7)}$$

due to

$$\int \mathrm{d}\Gamma\, x(\Gamma)\, y(\Gamma) = \int \mathrm{d}\Gamma\, y^2(\Gamma), \qquad \text{(A8)}$$

which is valid by constructions (A3–A6). Eq. (A2) follows from (A7). To get from here Eq. (15) one identifies: $x(\Gamma) = \sqrt{\mathcal{P}(\Gamma)}$, $b = \sqrt{\mathcal{P}(\Gamma)}\, e^{-\beta w(\Gamma)}$, $a = \sqrt{\mathcal{P}(\Gamma)}\, (f(\Gamma) - \langle f \rangle)$.

APPENDIX B: DERIVATION OF EQ. (91).

Let $f(x)$ is a smooth function, $\{x_i\}_{i=1}^n$ are n points, and

$$\bar{x} = \sum_{k=1}^n \lambda_k x_k, \quad \lambda_k \geq 0, \quad \sum_{k=1}^n \lambda_k = 1. \qquad \text{(B1)}$$

Apply the incomplete Taylor expansion to $f(x_i)$:

$$f(x_i) = f(\bar{x}) + f'(\bar{x})\ (x_i - \bar{x}) + \frac{f''(\xi_i)}{2}\ (x_i - \bar{x})^2, \qquad \text{(B2)}$$

where ξ_i lies between x_i and $\bar{x}$. Denote by $x_{\max}$ and $x_{\min}$ the maximal and the minimal numbers among x_i. This implies $x_{\max} \geq \xi_k \geq x_{\min}$. Now assume that $f''(x)$ is monotonically decaying:

$$f''(x_{\max}) \geq f''(\xi_i) \geq f''(x_{\min}). \qquad \text{(B3)}$$

Using (B2,B3) one has

$$\sum_{k=1}^n \lambda_k f(x_k) - f(\bar{x}) = \frac{1}{2} \sum_{k=1}^n f''(\xi_k) \lambda_k (x_k - \bar{x})^2, \qquad \text{(B4)}$$

$$\frac{f''(x_{\min})}{2} \sum_{k=1}^n \lambda_k (x_k - \bar{x})^2 \geq \sum_{k=1}^n \lambda_k f(x_k) - f(\bar{x}) \geq \frac{f''(x_{\max})}{2} \sum_{k=1}^n \lambda_k (x_k - \bar{x})^2. \qquad \text{(B5)}$$

To derive Eq. (91), start from (90), apply to it Eq. (B5) with a convex function $f(x) = e^{-\beta x}$, $\beta = 1/T > 0$, and identify $x_i = \varepsilon_i$, $\lambda_k = C_{kl}$. The desired Eq. (91) is recovered upon the summation over l.

APPENDIX C

Let $\mathrm{tr}(\hat{A}^2\hat{\rho}) = [\mathrm{tr}(\hat{A}\hat{\rho})]^2$ for some hermitian operator $\hat{A}$ and density matrix $\hat{\rho}$. In the main text we called such operators dispersionless with respect to the ensemble described by the density matrix $\hat{\rho}$.

In Cauchy inequality $|\mathrm{tr}(\hat{A}\hat{B})|^2 \leq \mathrm{tr}(\hat{A}\hat{A}^\dagger)\,\mathrm{tr}(\hat{B}\hat{B}^\dagger)$, which is valid for any operators $\hat{A}$ and $\hat{B}$, the equality is realized for $\hat{A} = \alpha\hat{B}^\dagger$, where α is a number. Thus

$$\left[\mathrm{tr}(\hat{A}\,\sqrt{\hat{\rho}}\,\sqrt{\hat{\rho}})\right]^2 = \mathrm{tr}(\hat{A}^2\hat{\rho})\,\mathrm{tr}(\hat{\rho}) \tag{C1}$$

implies

$$\hat{A}\sqrt{\hat{\rho}} = \alpha\,\sqrt{\hat{\rho}}. \tag{C2}$$

Now apply the eigenresolution $\sqrt{\hat{\rho}} = \sum_{k=1}^{n}\sqrt{p_k}\,|\varepsilon_k\rangle\langle\varepsilon_k|$ into (C2) and multiply it from the right by $|p_m\rangle$ to obtain:

$$\sqrt{p_m}\,\hat{A}\,|p_m\rangle = \alpha\,\sqrt{p_m}\,|p_m\rangle. \tag{C3}$$

It is seen that either only one among the eigenvalues p_k's is non-zero and then the corresponding eigenvector is also an eigenvector for $\hat{A}$, or, more generally, that $\hat{A}$ acts as $\propto \hat{1}$ in the Hilbert space formed by eigenvectors of $\hat{\rho}$ corresponding to its non-zero eigenvalues. In both cases the measurement of $\hat{A}$ on the state $\hat{\rho}$ always produces definite results.

Thus any operator $\hat{A}$ that is dispersionless on the density matrix $\hat{\rho}$ has to have the following block-diagonal matrix representation

$$\hat{A} = \begin{pmatrix} \alpha\,\hat{1}_{k\times k} & 0 \\ 0 & \hat{B} \end{pmatrix}, \tag{C4}$$

where α is a real number, $\hat{1}_{k\times k}$ is $k\times k$ unity matrix in the k-dimensional Hilbert space formed by eigenvectors corresponding to non-zero eigenvalues of $\hat{\rho}$, and finally $\hat{B}$ is an arbitrary $(n-k)\times(n-k)$ hermitian matrix. It has $(n-k)^2$ free parameters, and another free parameter of $\hat{A}$ is coming with the real number α. Thus, $\hat{A}$ has

$$(n-k)^2 + 1,$$

free parameters.

Note finally that various operators that are dispersionless on a pure density matrix need not be mutually commuting. As one of the simplest examples take

$$\hat{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ \hat{D} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & \epsilon \end{pmatrix},\ |\psi\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

where ϵ is real. As seen, $\hat{C}|\psi\rangle = |\psi\rangle$ and $\hat{D}|\psi\rangle = \epsilon\,|\psi\rangle$, but $[\hat{C},\hat{D}] \neq 0$.

APPENDIX D: RELATION BETWEEN POVMS AND PROJECTIVE MEASUREMENTS.

We outline how a POVM given by (43, 44, 45) can be connected with the usual projective measurements.

Let the system $\mathcal{S}$ interact with another auxiliary system $\mathcal{G}$. The initial states of $\mathcal{S}$ and $\mathcal{G}$ are, respectively, $\hat{\rho}$ (living in a n-dimensional Hilbert space $\mathcal{H}$) and $|g_1\rangle\langle g_1|$ living in a N-dimensional Hilbert space $\mathcal{H}'$. The initial state of the overall system $\mathcal{S}+\mathcal{G}$ is $\hat{\rho}\otimes|g_1\rangle\langle g_1|$.

Select a fixed orthonormal base $\{|u_k\rangle\}_{k=1}^{n}$ in $\mathcal{H}$. Let as well $|g_1\rangle$ be a member of an orthonormal base $\{|g_\alpha\rangle\}_{\alpha=1}^{N}$ in $\mathcal{H}'$. Assume that the above interaction is chosen such that the corresponding unitary evolution operator $\hat{\mathcal{U}}$ in the composite Hilbert space $\mathcal{H}\otimes\mathcal{H}'$ results in

$$\hat{\mathcal{U}}\,|u_k\rangle\otimes|g_1\rangle = \sum_{\alpha=1}^{N}\hat{G}_\alpha|u_k\rangle\otimes|g_\alpha\rangle, \tag{D1}$$

where $\hat{G}_\alpha$ defined by (43) are operators living in the Hilbert space $\mathcal{H}$.

One notes that due to (43)

$$\langle g_1|\otimes\langle u_l|\,\hat{\mathcal{U}}\,|u_k\rangle\otimes|g_1\rangle = \delta_{kl}. \tag{D2}$$

This specification of $\hat{\mathcal{U}}$ is not yet complete. To complete the definition of $\hat{\mathcal{U}}$ in the composite Hilbert space $\mathcal{H}\otimes\mathcal{H}'$ one should define its action on $|u_k\rangle\otimes|g_\alpha\rangle$ for $\alpha = 2, ..., N$ in addition to (D1). This will suffice, since $|u_k\rangle\otimes|g_\alpha\rangle$ for $k = 1, ..., n$ and $\alpha = 1, ..., N$ is an orthonormal base in the composite Hilbert $\mathcal{H}\otimes\mathcal{H}'$. This completion is possible to do and one can do that in many different ways, because it amounts to completing the orthonormal set of vectors $\hat{G}_\alpha|u_k\rangle\otimes|g_\alpha\rangle$ to an orthonormal base in $\mathcal{H}\otimes\mathcal{H}'$. Then the columns (or equivalently the rows) of $\hat{\mathcal{U}}$ in the base $|u_k\rangle\otimes|g_\alpha\rangle$ will be a set of of Nn orthonormal vectors, which is equivalent for $\hat{\mathcal{U}}$ being a unitary matrix. For a given unitary matrix $\hat{\mathcal{U}}$ there is a hermitian operator $\hat{H}_{\mathrm{ov}}$ in $\mathcal{H}\otimes\mathcal{H}'$ such that $\hat{\mathcal{U}} = \exp\left[\frac{it}{\hbar}\hat{H}_{\mathrm{ov}}\right]$ with some time-parameter t. Thus, $\hat{H}_{\mathrm{ov}}$ can serve as a Hamiltonian realizing the needed interaction.

Once (D1) is valid for the fixed base $\{|u_k\rangle\}_{k=1}^{n}$, one gets for an arbitrary $\hat{\rho}$ in $\mathcal{H}$:

$$\hat{\mathcal{U}}\,\hat{\rho}\,\hat{\mathcal{U}}^\dagger = \sum_{\alpha=1}^{N}\hat{G}_\alpha\,\hat{\rho}\,\hat{G}_\alpha^\dagger\otimes|g_\alpha\rangle\langle g_\alpha|. \tag{D3}$$

To complete the construction, one now measures in $\mathcal{H}'$ any hermitian operator with a non-degenerate spectrum having the base $\{|g_\alpha\rangle\}_{\alpha=1}^{N}$ as its eigenbase. Then the POVM (43, 44, 45) accounts for what is happening — after the interaction and after the selective measurement — with the initial ensemble described by $\hat{\rho}$.

APPENDIX E: DERIVATION OF EQS. (103, 104).

One notes from (21, 28):

$$\hat{U}_\tau^\dagger\,\hat{\rho}\,\hat{U}_\tau = \frac{\exp\left[-\beta\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau\right]}{Z} = \frac{e^{-\beta\hat{\Omega}-\beta\hat{H}}}{Z}, \qquad \text{(E1)}$$

where we used the definition $\hat{\Omega} = \hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau - \hat{H}$ of $\hat{\Omega}$.

Note the standard equality

$$e^{-\beta\hat{\Omega}-\beta\hat{H}} = \overrightarrow{\exp}\left[-\int_0^\beta \mathrm{d}s\, e^{-s\hat{H}}\,\hat{\Omega}\, e^{s\hat{H}}\right]\, e^{-\beta\hat{H}}, \qquad \text{(E2)}$$

where $\overrightarrow{\exp}$ means time-antiordered exponent. The easiest way to derive Eq. (E2) is to differentiate both sides of it over β, and note that they both satisfy to the same first-order differential equation with the same boundary condition at $\beta = 0$.

One now gets

$$\hat{U}_\tau^\dagger\,\hat{\rho}\,\hat{U}_\tau = \overrightarrow{\exp}\left[-\int_0^\beta \mathrm{d}s\, e^{-s\hat{H}}\hat{\Omega}\, e^{s\hat{H}}\right]\,\hat{\rho}. \qquad \text{(E3)}$$

Tracing out both sides, one finally obtains (103, 104).

The simplest way to derive Peierls-Bogoliubov inequality from (103, 104) is to note the well-known extremal feature of the free energy:

$$\begin{aligned} &-T\ln \mathrm{tr} e^{-\beta\hat{H}-\beta\hat{\Omega}} \\ &= min\left\{\mathrm{tr}[\hat{\rho}\,(\hat{H}+\hat{\Omega})] + T\mathrm{tr}(\hat{\rho}\ln\hat{\rho})\right\}, \end{aligned}$$

where the minimization is taken over all possible density matrices. This can alternatively be written as

$$\begin{aligned} &\mathrm{tr}\, e^{-\beta\hat{H}-\beta\hat{\Omega}} \\ &= max \exp\left\{-\beta\mathrm{tr}[\hat{\rho}\,(\hat{H}+\hat{\Omega})] - \mathrm{tr}(\hat{\rho}\ln\hat{\rho})\right\}. \end{aligned} \qquad \text{(E4)}$$

The desired inequality is got by putting the particular density matrix $\hat{\rho} = e^{-\beta\hat{H}}/Z$ in the RHS of (E4).

APPENDIX F: PROOF OF HORN'S THEOREM.

We intend to prove that given two vectors $x^{\mathrm{T}} = (x_1 \geq ... \geq x_n)$ and $y^{\mathrm{T}} = (y_1 \geq ... \geq y_n)$ with the following majorization relation

$$x \prec y \qquad \text{(F1)}$$

there is a real orthogonal matrix $O = (O_{ij})$ such that

$$x_i = \sum_j O_{ij}^2 y_j \;\Leftrightarrow\; x = \mathrm{diag}\left[\, O\,\mathrm{diag}[y]\, O^{\mathrm{T}}\,\right]. \qquad \text{(F2)}$$

Here diag$[y]$ means the $n\times n$ diagonal matrix formed by the vector y, diag$\left[\, O\,\mathrm{diag}[y]\, O^{\mathrm{T}}\,\right]$ is the vector formed by diagonal elements of the matrix $O\,\mathrm{diag}[y]\, O^{\mathrm{T}}$, and O^{T} means transposition: $(O^{\mathrm{T}})_{kl} = O_{lk}$. Note that for any orthogonal matrix O_{ij}, the matrix O_{ij}^2 is always double-stochastic: $\sum_i O_{ij}^2 = \sum_j O_{ij}^2 = 1$, while the converse is not true [56].

For a given orthogonal matrix O_{ij} there are many unitary matrices Q_{ij} such that $O_{ij}^2 = |Q_{ij}|^2$; e.g., $Q_{ij} = e^{i\phi_j}O_{ij}$, where ϕ_j are arbitrary phases.

The proof is adopted from Ref. [59] and will be realized in two steps.

1. First step.

First, one shows that (F1) implies

$$x = T_1\, T_2 ... T_{n-1}\, y, \qquad \text{(F3)}$$

where matrices T are so-called T-transform defined as follows. Any T-transform $T(m,l;t)$ has three parameters: $m < l$ and t, where m and l are natural numbers between zero and n, and where $0 < t < 1$. Its action on any vector y, $y^{\mathrm{T}} = (y_1, ..., y_n)$, is defined as

$$z = T(m,l;t)y, \qquad \text{(F4)}$$

where

$$\begin{aligned} z^{\mathrm{T}} = (y_1, .., y_{m-1}, ty_m + (1-t)y_l, y_{m+1}, ..., \\ y_{l-1}, (1-t)y_m + ty_l, y_{l+1}, ..., y_n) \end{aligned} \qquad \text{(F5)}$$

To get the matrix of $T(m,l;t)$ starting from $n\times n$ unity matrix $\hat{1}$, one proceeds as follows:

$$\begin{aligned} &(\hat{1})_{mm} = 1 \;\rightarrow\; (\,T(m,l;t)\,)_{mm} = t, \\ &(\hat{1})_{ll} = 1 \;\rightarrow\; (\,T(m,l;t)\,)_{ll} = t, \\ &(\hat{1})_{ml} = 0 \;\rightarrow\; (\,T(m,l;t)\,)_{ml} = 1-t, \\ &(\hat{1})_{lm} = 0 \;\rightarrow\; (\,T(m,l;t)\,)_{lm} = 1-t, \end{aligned} \qquad \text{(F6)}$$

while all other elements of the unity matrix are left unchanged.

Eq. (F3) can now be proven by induction. It is obvious for $n = 2$. Assume it holds for $n-1$. As seen from (F1, 73, 74), one has $y_n \leq x_1 \leq y_1$, so there exists an index k that

$$y_k \leq x_1 \leq y_1. \qquad \text{(F7)}$$

This implies

$$x_1 = ty_1 + (1-t)y_k, \qquad \text{(F8)}$$

for some $0 \leq t \leq 1$. Define a T-transform $T(1,k;t)$ via

$$\begin{pmatrix} x_1 \\ \\ \bar{y} \end{pmatrix} = T(1,k;t)\, y, \qquad \text{(F9)}$$

where

$$\bar{y}^{\mathrm{T}} = (y_2, ..., y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, .., y_n). \quad \text{(F10)}$$

It is straighforward to show that

$$\bar{y}^{\mathrm{T}} \succ (x_2, .., x_n). \quad \text{(F11)}$$

Since we assumed (F1) $\Rightarrow$ (F3) is valid for $n-1$, there is a product of T-transforms such that $(x_2, .., x_n)^{\mathrm{T}} = T_2 ... T_{n-1}\, \bar{y}$.

Thus the implication (F1) $\Rightarrow$ (F3) is proven by induction.

2. Second step.

Let us finally prove the implication (F1) $\Rightarrow$ (F2). Note that for any T-transform $T(m,l;t)$ one can associate an orthogonal matrix $V(m,l;t)$ by reshaping (F6) as follows:

$$\begin{aligned}
&(\hat{1})_{mm} = 1 \;\rightarrow\; (\, V(m,l;t)\,)_{mm} = \sqrt{t},\\
&(\hat{1})_{ll} = 1 \;\rightarrow\; (\, V(m,l;t)\,)_{ll} = \sqrt{t},\\
&(\hat{1})_{ml} = 0 \;\rightarrow\; (\, V(m,l;t)\,)_{ml} = -\sqrt{1-t},\\
&(\hat{1})_{lm} = 0 \;\rightarrow\; (\, V(m,l;t)\,)_{lm} = \sqrt{1-t}.
\end{aligned} \quad \text{(F12)}$$

Then (F4) is equivalent to

$$z = \mathrm{diag}\left[\, V(m,l;t)\ \mathrm{diag}[y]\ V^{\mathrm{T}}(m,l;t)\,\right]. \quad \text{(F13)}$$

To prove the implication (F1) $\Rightarrow$ (F2) one again proceeds by induction. It is obviously valid for $n=2$. One assumes its validity for $n-1$. This means (F11) can be re-written as:

$$(x_2, ..., x_n)^{\mathrm{T}} = \mathrm{diag}\left[\, \tilde{V}\ \mathrm{diag}[\bar{y}]\ \tilde{V}^{\mathrm{T}}\,\right], \quad \text{(F14)}$$

where $\tilde{V}$ is some orthogonal matrix. To complete the proof, define an orthogonal matrix

$$O = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{V} \end{pmatrix} V, \quad \text{(F15)}$$

where the matrix V corresponds to the T-transform T defined in (F9) (via the correspondence described in (F12)), and rewrite (F11, F9) as

$$x = \mathrm{diag}\left[\, O\ \mathrm{diag}[\bar{y}]\ O^{\mathrm{T}}\,\right]. \quad \text{(F16)}$$

This proves the implication (F1) $\Rightarrow$ (F2).

3. An example.

Let us realize explicitly the above construction for an example:

$$x = (0,0,0), \qquad y = (2,1,-3). \quad \text{(F17)}$$

For the index k and for the parameter t mentioned before (F7) one has

$$k = 3, \qquad t = \frac{3}{5}. \quad \text{(F18)}$$

Analogously:

$$\begin{pmatrix} 0 \\ \\ 0 \\ \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{3}{5} & 0 & \frac{2}{5} \\ \\ 0 & 1 & 0 \\ \\ \frac{2}{5} & 0 & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 2 \\ \\ 1 \\ \\ -3 \end{pmatrix} \quad \text{(F19)}$$

realizes the relation between x and y via T-transforms. Finally the orthogonal matrix O in (F16) reads for the present example:

$$O = \begin{pmatrix} \sqrt{\frac{3}{5}} & 0 & -\sqrt{\frac{2}{5}} \\ \\ -\sqrt{\frac{1}{5}} & \sqrt{\frac{1}{2}} & -\sqrt{\frac{3}{10}} \\ \\ \sqrt{\frac{1}{5}} & \sqrt{\frac{1}{2}} & \sqrt{\frac{3}{10}} \end{pmatrix}. \quad \text{(F20)}$$

APPENDIX G: DERIVATION OF EQ. (60).

Here we find the maximum of

$$\langle w^2 \rangle = \sum_{\alpha=1}^{N} \lambda_\alpha w_\alpha^2 = \sum_{\alpha=1}^{N} \frac{\widetilde{\langle \psi_\alpha | \hat{\Omega} | \psi_\alpha \rangle}^2}{\widetilde{\langle \psi_\alpha | \psi_\alpha \rangle}}, \quad \text{(G1)}$$

where the maximization is taken over all possible decompositions (34) of the mixed state $\hat{\rho}$ into pure states. Using (54, 56) one writes equivalently

$$\langle w^2 \rangle = \sum_{\alpha=1}^{N} \frac{\left[\, \mathrm{tr}\,(\hat{\rho}^{1/2} \hat{\Omega} \hat{\rho}^{1/2}\, \hat{\Pi}_\alpha)\,\right]^2}{\mathrm{tr}\, \hat{\rho} \hat{\Pi}_\alpha}, \quad \text{(G2)}$$

where

$$\hat{\Pi}_\alpha = |\pi_\alpha\rangle\langle\pi_\alpha|. \quad \text{(G3)}$$

The maximization in (G2) is taken over all decompositions of unity

$$\sum_{\alpha=1}^{N} \hat{\Pi}_\alpha = \hat{1}, \quad \text{(G4)}$$

where operators $\hat{\Pi}_\alpha$ live in the n-dimensional Hilbert space $\mathcal{H}$.

The general idea of the following method was adopted from [62]. Introduce an operator $\hat{X}$ via

$$\hat{\rho}^{1/2} \hat{\Omega} \hat{\rho}^{1/2} = \Re(\hat{\rho}\hat{X}) \equiv \frac{1}{2}\left(\hat{\rho}\hat{X} + \hat{X}\hat{\rho}\right), \quad \text{(G5)}$$

then

$$\mathrm{tr}(\hat{\rho}^{1/2}\hat{\Omega}\hat{\rho}^{1/2}\hat{\Pi}_\alpha) = \Re\,\mathrm{tr}(\hat{\Pi}_\alpha\hat{\rho}\hat{X}). \tag{G6}$$

Recall Cauchy inequality

$$|\mathrm{tr}(\hat{A}\hat{B})|^2 \le \mathrm{tr}(\hat{A}\hat{A}^\dagger)\,\mathrm{tr}(\hat{B}\hat{B}^\dagger), \tag{G7}$$

which is valid for any operators $\hat{A}$ and $\hat{B}$, with the equality being realized for

$$\hat{A} = \alpha\hat{B}^\dagger, \tag{G8}$$

where α is a number.

Applying (G5, G7):

$$\left[\mathrm{tr}\,(\hat{\rho}^{1/2}\hat{\Omega}\hat{\rho}^{1/2}\,\hat{\Pi}_\alpha)\right]^2 \le |\mathrm{tr}(\hat{\Pi}_\alpha\hat{\rho}\hat{X})|^2 \tag{G9}$$

$$\equiv |\mathrm{tr}(\hat{\Pi}_\alpha^{1/2}\,\hat{\rho}^{1/2}\,\hat{\rho}^{1/2}\,\hat{X}\,\hat{\Pi}_\alpha^{1/2})|^2$$

$$\le \mathrm{tr}(\hat{\Pi}_\alpha\,\hat{\rho})\;\mathrm{tr}(\hat{\rho}\,\hat{X}\,\hat{\Pi}_\alpha\,\hat{X}), \tag{G10}$$

one gets for (G2, G3):

$$\sum_{\alpha=1}^{N}\frac{\left[\mathrm{tr}\,(\hat{\rho}^{1/2}\hat{\Omega}\hat{\rho}^{1/2}\,\hat{\Pi}_\alpha)\right]^2}{\mathrm{tr}\,\hat{\rho}\hat{\Pi}_\alpha} \le \sum_{\alpha=1}^{N}\mathrm{tr}(\hat{\rho}\,\hat{X}\,\hat{\Pi}_\alpha\,\hat{X}) \tag{G11}$$

$$= \mathrm{tr}(\hat{\rho}\hat{X}^2) = \mathrm{tr}(\hat{\rho}^{1/2}\hat{\Omega}\hat{\rho}^{1/2}\hat{X}) \tag{G12}$$

Eq. (G9) is realized as equality for

$$\Im\,\mathrm{tr}(\hat{\Pi}_\alpha\hat{\rho}\hat{X}) = 0, \tag{G13}$$

while the Cauchy inequality (G11) becomes equality for

$$\hat{\rho}^{1/2}\,\hat{X}|\pi_\alpha\rangle = a_\alpha\hat{\rho}^{1/2}|\pi_\alpha\rangle, \tag{G14}$$

where a_α are some numbers.

Both conditions (G13, G14) are realized simultaneously by taking $\{|\pi_\alpha\rangle\}_{\alpha=1}^n$ and $\{a_\alpha\}_{\alpha=1}^n$ as, respectively, eigenvectors and eigenvalues of the hermitian operator $\hat{X}$. The representation (62) for $\hat{X}$ follows from (G5). The desired equation (60) is seen from (G11, G12).

APPENDIX H: DERIVATION OF EQ. (80).

Here we calculate the average $\{\langle w^2\rangle\}_{\rm av}$ of $\langle w^2\rangle$, given by (G1), over the measure (80). Using Eq. (49) it is straightforward to see that all the terms in the summation in the RHS of (G1) produce the same average. Thus,

$$\frac{\{\langle w^2\rangle\}_{\rm av}}{N} = \frac{\int\mathcal{D}M\,\delta\left[\sum_{i=1}^N|M_i|^2-1\right]\phi\{M_i\}}{\int\mathcal{D}M\,\delta\left[\sum_{i=1}^N|M_i|^2-1\right]}, \tag{H1}$$

where we denoted

$$\mathcal{D}M = \prod_{i=1}^{N}\mathrm{d}\Re M_i\,\mathrm{d}\Im M_i \tag{H2}$$

and where one notes from (49)

$$\phi\{M_i\} = \left|\sum_{j,k=1}^{n} M_jM_k^*\sqrt{p_jp_k}\langle\varepsilon_k|\hat{\Omega}|\varepsilon_j\rangle\right|^2. \tag{H3}$$

Passing to polar coordinates

$$\int\mathcal{D}M = \int_0^{2\pi}\prod_{i=1}^{N}\mathrm{d}\varphi_i\int_0^\infty\prod_{i=1}^{N}|M_i|\,\mathrm{d}|M_i| \tag{H4}$$

one gets

$$\frac{\{\langle w^2\rangle\}_{\rm av}}{N} = \sum_{j,k=1}^{n}p_jp_k\langle\varepsilon_j|\hat{\Omega}|\varepsilon_j\rangle\langle\varepsilon_k|\hat{\Omega}|\varepsilon_k\rangle\,\frac{I_{ij}}{I_0}, \tag{H5}$$

where

$$I_{jk} = \int_0^\infty\prod_{i=1}^{N}\mathrm{d}z_i\;\;\delta\left[\sum_{i=1}^{N}z_i-1\right]\;\frac{z_jz_k}{\sum_{l=1}^n p_lz_l}, \tag{H6}$$

$$I_0 = \int_0^\infty\prod_{i=1}^{N}\mathrm{d}z_i\;\;\delta\left[\sum_{i=1}^{N}z_i-1\right]. \tag{H7}$$

These integrals are calculated for $j,k=1,..,n$ by the same method. For example,

$$e^{-r}r^NI_{jj} = \int_0^\infty\prod_{i=1}^{N}\mathrm{d}y_i\;\;\delta\left[\sum_{i=1}^{N}y_i-r\right]\;\frac{y_j^2\,e^{-r}}{\sum_{l=1}^n p_ly_l},$$

$$\Gamma(N+1)I_{jj} = \int_0^\infty\prod_{i=1}^{N}\mathrm{d}y_i\,y_j^2\;\;\frac{e^{-\sum_{i=1}^N y_i}}{\sum_{l=1}^n p_ly_l}, \tag{H8}$$

$$= \int_0^\infty\prod_{l=1}^{n}\mathrm{d}y_l\,y_j^2\;\;\frac{e^{-\sum_{l=1}^n y_l}}{\sum_{l=1}^n p_ly_l}, \tag{H9}$$

$$= \int_0^\infty\mathrm{d}s\int_0^\infty\prod_{i=1}^{N}\mathrm{d}y_i\,y_j^2\;\;e^{-\sum_{l=1}^n y_l(sp_l+1)}$$

where when passing from (H8) from (H9) we changed the integration variable, $z_i = y_i/r$ and integrated over r from 0 to ∞.

Further calculations are straightforward and lead to (80). For dealing with this equation the following formula is useful:

$$\int_0^\infty\mathrm{d}s\prod_{k=1}^{n}\frac{1}{\theta_m+s} = \sum_{k=1}^{n}\ln\theta_k\prod_{l\ne k}^{n}\frac{1}{\theta_l-\theta_k}, \tag{H10}$$

where θ_k's are some positive numbers.

APPENDIX I

Here we give an example of the situation discussed around (94, 95). The effect announced there does exist neither for two-level systems —simply because for a

2×2 traceless matrix $\hat{\Omega}$ a zero eigenvalue implies $\hat{\Omega}=0$,—nor for three level systems. This last fact requires some calculations which will be omitted.

The simplest situation which supports the effect is thus a four-level system. The following example was inspired by [63]. Consider a four-level system with Hamiltonian:

$$\hat{H}=\begin{pmatrix}\hat{A} & 0\\ 0 & \hat{B}\end{pmatrix}, \tag{I1}$$

where $\hat{A}$ and $\hat{B}$ are 2×2 matrices:

$$\hat{A}=\begin{pmatrix}a & b\\ b & d\end{pmatrix}, \quad \hat{B}=\begin{pmatrix}a & b\\ b & c\end{pmatrix}, \tag{I2}$$

with a, b, c and d being some real numbers.

Assume that the unitary operator $\hat{U}_\tau$ is given as an exchange interaction:

$$\hat{U}_\tau=\begin{pmatrix}0 & \hat{1}\\ \hat{1} & 0\end{pmatrix}, \tag{I3}$$

where $\hat{1}$ is the 2×2 unit matrix.

The Hamiltonian $\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau$ in the Heisenberg representation at time τ then reads

$$\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau=\begin{pmatrix}\hat{B} & 0\\ 0 & \hat{A}\end{pmatrix}. \tag{I4}$$

It is seen from the block-diagonal form of $\hat{H}$ and $\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau$ that they both have the same eigenvalues.

As follows from (I1, I2, I4), the matrix $\hat{\Omega}$,

$$\hat{\Omega}=\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & c-d & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & d-c\end{pmatrix}, \tag{I5}$$

has a doubly degenerate eigenvalue equal to zero, and the corresponding eigenvectors can be taken as

$$|0_1\rangle=\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad |0_2\rangle=\begin{pmatrix}0\\0\\1\\0\end{pmatrix}. \tag{I6}$$

It is now obvious that though

$$\left\langle 0_1\left|\left[\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau\right]^m\right|0_1\right\rangle-\langle 0_1|\hat{H}^m|0_1\rangle=0, \quad \text{for} \quad m=1,2,$$

one still has

$$\left\langle 0_1\left|\left[\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau\right]^3\right|0_1\right\rangle-\langle 0_1|\hat{H}^3|0_1\rangle=b^2(c-d)\neq0,$$

$$\left\langle 0_1\left|\left[\hat{U}_\tau^\dagger\,\hat{H}\,\hat{U}_\tau\right]^4\right|0_1\right\rangle-\langle 0_1|\hat{H}^4|0_1\rangle = 2ab^2(c-d)+b^2(c^2-d^2)\neq0. \tag{I7}$$

These relations were used in (96).